DEVELOPING CORRECTION FACTORS FOR REFERENCE FLEET. THE CASE OF ATLANTA I/M PROGRAM EVALUATION.

A Dissertation
Presented to
The Academic Faculty

by

Olga Kemenova

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of School of Civil and Environmental Engineering

Georgia Institute of Technology
May, 2020

DEVELOPING CORRECTION FACTORS FOR REFERENCE FLEET. THE CASE OF ATLANTA I/M PROGRAM EVALUATION.

Approved by:

Dr. Michael O. Rodgers, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Catherine L. Ross
School of City and Regional Planning/Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Randall L. Guensler
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Wayne Daley
GTRI/ATAS
Georgia Institute of Technology

Dr. Michael P. Hunter
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: January 10th, 2020

To my family for their patience and persistent support.

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude to the members of my committee for their guidance and support. I am especially thankful for Dr. Michael Rodgers, the chair of my committee, who believed in my ability to complete this work even when I myself had doubts. I am also grateful to Dr. Rodgers and Dr. Leisha Davis, who almost 16 years ago, hired me as Dr. Davis's replacement to perform Inspection and Maintenance Program Evaluation.

This research would not be possible without longstanding support from the Georgia Department of Natural Resources for the CAFÉ program.

Special thanks to the Georgia Tech Research Institute for providing fundings for my studies and the completion of the dissertation research.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS		**vii**
LIST OF TABLES		**xii**
LIST OF FIGURES		**xiii**
LIST OF SYMBOLS AND ABBREVIATIONS		**xvi**
SUMMARY		**xviii**
CHAPTER 1.	**Introduction**	**1**
CHAPTER 2.	**Background**	**5**
2.1	**Vehicle Inspection/Maintenance Programs**	**5**
2.2	**Inspection/Maintenance in Atlanta**	**6**
2.3	**Existing Enhanced I/M Program Evaluation Methods**	**8**
2.3.1	Evaluation based on Emission Inspection Records:	9
2.3.2	Evaluation based on Roadside Emission Inspections:	12
2.3.3	Evaluation Methods based on Data from Onroad Vehicles:	14
CHAPTER 3.	**Research Design Components**	**20**
3.1	**Description of Research Setting**	**20**
3.2	**On Road Emission Data Collection**	**21**
3.3	**Research Objective and Justification. General Reference Method. Case of Atlanta I/M Evaluation**	**26**
3.4	**Research Questions**	**30**
3.5	**Data Description**	**31**
3.6	**Methodology and Analysis Plan**	**33**
CHAPTER 4.	**Analysis**	**39**
4.1	**Examining Factors Affecting Vehicle Emissions Measurements**	**39**
4.1.1	Vehicles' Age, Fleet Composition	39
4.1.2	Income, Fleet "Make-Model" Distribution	42
4.1.3	Ambient Temperature	46
4.1.4	Emission Print and Driving Conditions	50
4.1.5	Comparing Emissions within Same VSP Range	56
4.1.6	VSP Distributions within Experimental Sites and Counties	60
4.2	**Null Experiment: Comparing Fulton Light Duty Fleet with Gwinnett Light Duty Fleet**	**62**
4.3	**Calculating I/M Effectiveness while Controlling for VSP**	**65**
4.4	**Calculating I/M Effectiveness while Controlling by VSP bins defined by MOVES Model**	**73**
CHAPTER 5.	**Conclusions**	**75**
5.1	**Overall Results**	**75**

5.2 Main Contributions to the Field **78**

APPENDIX A. Biennial evaluation of the emissions reduction effectiveness of the Atlanta vehicle Inspection and maintenance program for 2009-2010 **81**
A.1 Introduction **82**
A.1.1 Overview 82
A.1.2 Vehicle Inspection/Maintenance Programs: 85
A.1.3 Inspection/Maintenance In Atlanta: 86
A.2 Enhanced I/M Program Evaluation Methods **88**
A.2.1 Emissions Inspection Records: 89
A.2.2 Roadside Emission Inspections: 92
A.2.3 Remote Sensing Data from Onroad Vehicles: 93
A.3 I/M Program Evaluation Components **98**
A.3.1 On Road Emissions Data: 99
A.3.2 Predicted Emission Factors: 103
A.3.3 Evaluation Algorithm: 103
A.4 Analysis **104**
A.4.1 Data Overview: 105
A.4.2 Validity of Fleet Comparisons: 106
A.4.3 Reference Method Results: 108
A.5 Discussion **112**
A.6 Comparing Results with Previous Reviews **113**

APPENDIX B. Biennial evaluation of the emissions reduction effectiveness of the atlanta vehicle Inspection and maintenance program for 2007-2008 **117**
B.1 Introduction **118**
B.1.1 Overview: 118
B.1.2 Vehicle Inspection/Maintenance Programs: 121
B.1.3 Inspection/Maintenance In Atlanta: 122
B.2 Enhanced I/M program Evaluation Methods **124**
B.2.1 Emissions Inspection Records: 125
B.2.2 Roadside Emission Inspections: 128
B.2.3 Remote Sensing Data from Onroad Vehicles: 129
B.3 I/M Program Evaluation Components **134**
B.3.1 On Road Emissions Data: 134
B.3.2 Predicted Emission Factors: 138
B.3.3 Evaluation Algorithm: 138
B.4 Analysis **139**
B.4.1 Data Overview: 140
B.4.2 Validity of Fleet Comparisons: 141
B.4.3 Reference Method Results: 143
B.5 Discussion **147**
B.6 Comparing Results with Previous Reviews **148**

APPENDIX C. Biennial evaluation of the emissions reduction effectiveness of the atlanta vehicle Inspection and maintenance program for 2005-2006 **152**
C.1 Introduction **153**

C.1.1 Overview: 153
C.1.2 Vehicle Inspection/Maintenance Programs: 156
C.1.3 Inspection/Maintenance in Atlanta: 157
C.2 Enhanced I/M Programs Evaluation Methods 159
C.2.1 Emissions Inspection Records: 160
C.2.2 Roadside Emission Inspections: 163
C.2.3 Remote Sensing Data from Onroad Vehicles: 164
C.3 I/M Program Evaluation Components 169
C.3.1 On-Road Emissions Data: 169
C.3.2 Predicted Emission Factors: 173
C.3.3 Evaluation Algorithm: 173
C.4 Analysis 175
C.4.1 Data Overview: 176
C.4.2 Validity of Fleet Comparisons: 177
C.4.3 Reference Method Results: 179
C.5 Discussion 183
C.6 Comparing Results with Previous Reviews 185

APPENDIX D. Biennial evaluation of vehicle inspection/maintenance program using on-road emission data for 2003-2004 188
D.1 Introduction 189
D.1.1 Vehicle Inspection/Maintenance Programs 192
D.1.2 Inspection/Maintenance In Atlanta 193
D.2 Enhanced I/M Programs Evaluations 195
D.2.1. Emissions Inspection Records 196
D.2.2 Roadside Emission Inspections 199
D.2.3 Remote Sensing Data from Onroad Vehicles 200
D.3 I/M Program Evaluation Components 205
D.3.1 Onroad Emissions Data 205
D.3.2 Predicted Emission Factors 209
D.3.3 Evaluation Algorithm 209
D.4 Analysis 210
D.4.1 Data Overview 211
D.4.2 Validity of Fleet Comparisons 212
D.4.3 Reference Method Results 214
D.5 Discussion 219
D.6 Comparing Results with Previous Reviews. 220

APPENDIX E. Biennial evaluation of vehicle inspection/maintenance program using on-road emission data for 2001-2002 223
E.1 Introduction 224
E.1.1 Vehicle Inspection/Maintenance Programs 227
E.1.2 Inspection/Maintenance In Atlanta 228
E.2 Enhanced I/M Programs Evaluations 230
E.2.1 Emissions Inspection Records 231
E.2.2 Roadside Emission Inspections 234
E.2.3 Remote Sensing Data from Onroad Vehicles 235

E.3 I/M Program Evaluation components **240**
E.3.1 Onroad Emissions Data 240
E.3.2 Predicted Emission Factors 244
E.3.3 Evaluation Algorithm 244
E.4 Analysis **245**
E.4.1 Data Overview 246
E.4.2 Validity of Fleet Comparisons 247
E.4.3 Reference Method Results 249
E.5 Discussion **255**
E.6 Comparing results with previous review. **256**

BIBLIOGRAPHY **259**

LIST OF TABLES

Table 4.1	ANOVA for model year distribution by counties subject to I/M testing.	41
Table 4.2	ANOVA for CO differences across observed temperature blocks. Atlanta 13 counties. For 6 < VSP < 9 kW/Tonne.	48
Table 4.3	ANOVA for CO differences across observed temperature blocks. Augusta/Macon. For 6 < VSP < 9 kW/Tonne.	48
Table 4.4	ANOVA for CO and NOx mean levels among Atlanta 13 I/M counties.	51
Table 4.5	ANOVA for CO and NOx mean levels among Atlanta 13 I/M counties for VSP less than 10kW/Tonne	56
Table 4.6	ANOVA for CO and NOx mean levels among range of VSP bins	58
Table 4.7	ANOVA for VSP distribution within Atlanta 13 counties' vehicles	61
Table 4.8	Emission reductions associated with I/M program by VSP bins. 2010 measurement year.	72
Table 4.9	Emission reductions associated with I/M program by VSP bins. 2010 measurement year. LDV's 2002 model year and newer.	72
Table 4.10	Emission reductions associated with I/M program by MOVES VSP bins. 2010 measurement year. LDV's all model years.	74
Table A. 1	Effectiveness of Atlanta I/M Program and Fuel Program.	108
Table A.2	Biennial I/M Effectiveness Estimated for 1998 - 2010 Measurement Years.	115
Table B.1	Effectiveness of Atlanta I/M Program and Fuel Program.	143
Table B.2	I/M Effectiveness Estimated for 1998, 2000, 2002, 2004, 2006 and 2008 Measurement Years.	150
Table C.1	Effectiveness of Atlanta I/M Program and Fuel Program.	180
Table C.2	I/M Effectiveness Estimated for 1998, 2000, 2002, 2004 and 2006 Measurement Years.	187
Table D.1	Effectiveness of Atlanta I/M Program and Fuel Program.	215
Table D.	IM Effectiveness Estimated for 1998, 2000, 2002, and 2004 Measurement Year.	221
Table E.1	Effectiveness of Atlanta I/M Program and Fuel Program.	249
Table E.2	IM Effectiveness Estimated for 1998, 2000 and 2002 Measurement Year.	257

LIST OF FIGURES

Figure 3.1	2009- 2010 Remote sensing sites for the Continuous Atlanta Fleet Evaluation (CAFÉ) in the thirteen-county Atlanta vehicle inspection and maintenance (I/M) area and Augusta and Macon, Georgia	23
Figure 3.2	Mean CO emission levels for vehicles manufactured by different companies. Measured in Atlanta thirteen-county area in 2010.	29
Figure 3.3	Dataset Development.	35
Figure 3.4	Data Reduction.	35
Figure 3.5	Data Split.	36
Figure 3.6	Mean CO values across Atlanta thirteen-county I/M area light duty fleets. 2008 measurement year.	37
Figure 4.1	Distribution of observed Light Duty (LD) fleet among Atlanta 13 counties.	39
Figure 4.2	Model year distribution among Atlanta 13 I/M counties. 2010 measurement year.	40
Figure 4.3	Atlanta 13 counties fleet distribution by vehicle body type. 2010 measurement year.	42
Figure 4.4	Income distribution among Atlanta 13 I/M counties. Source: U.S. Census Bureau, 2011 American Community Survey	43
Figure 4.5	Model year distribution vs. income distribution among Atlanta 13 counties.	44
Figure 4.6	Atlanta 13 counties fleet distribution by manufacturer. 2010 measurement year.	45
Figure 4.7	Fleet distribution by manufacturer. Selected counties. 2010 measurement year.	46
Figure 4.8	Data collection by temperature block. Atlanta 13 counties area vs. Augusta/Macon. 2010 measurement year.	47
Figure 4.9	Mean NOx levels for Atlanta 13 counties and Augusta/Macon by temperature blocks. For $6 < VSP < 9$ kW/Tonne. 2010 Measurement Year.	49
Figure 4.10	Mean NOx levels for Atlanta 13 counties by temperature blocks. For $6 < VSP < 9$ kW/Tonne. Normalized to the same model year distribution. 2010 measurement year.	50
Figure 4.11	95% Confidence intervals for CO emission differences among Fulton and other counties, 2010 measurement year.	52
Figure 4.12	95% Confidence intervals for NOx emission differences among Fulton and other counties, 2010 measurement year.	52
Figure 4.13	Mean CO emissions across Atlanta 13 counties LD fleets for VSP levels less than 10kW/Tonne	54
Figure 4.15	Mean CO emissions for different VSP bins	57

Figure 4.16	VSP distribution in Atlanta 13 county and Augusta/Macon areas	59
Figure 4.17	VSP bin distributions in Atlanta and Augusta/Macon areas	60
Figure 4.18	Mean VSP levels observed at different measurement site locations	61
Figure 4.19	95% Confidence intervals for mean CO differences between Fulton and Gwinnet LD fleets comparing with difference predicted by MOVES for Atlanta 13 counties area with and without I/M program. Controlled for MY and VSP. 2010 measurement year. All LD vehicles.	64
Figure 4.20	95% Confidence intervals for mean NOx differences between Fulton and Gwinnet LD fleets comparing with difference predicted by MOVES for Atlanta 13 counties area with and without I/M program. Controlled for MY and VSP. 2010 measurement year. All LD vehicles.	65
Figure 4.21	Mean CO Emission reductions in percentiles associated with I/M program by VSP bins. 2010 measurement year. Controlled by MY but unweighted by number of cars within the bin.	67
Figure 4.22	Mean NOx Emission reductions in percentiles associated with I/M program by VSP bins. 2010 measurement year. Controlled by MY but unweighted by number of cars within the bin.	68
Figure 4.23	CO means by model year. All LD vehicles. Atlanta 13 counties inspected fleet vs. Augusta/Macon fleets. Controlled by MY distribution. Controlled by VSP: VSP < 10kW/Tonne. 2010 Measurement Year.	69
Figure 4.24	NOx means by model year. All LD vehicles. Atlanta 13 counties inspected fleet vs. Augusta/Macon fleets. Controlled by MY distribution. Controlled by VSP: VSP < 10kW/Tonne. 2010 Measurement Year.	70
Figure 4.25	Mean CO and NOx reductions in percentiles within MOVES VSP bins. Atlanta 13 counties - Augusta/Macon case. Controlled for MY but unweighted by number of cars within VSP bins. 2010 measurement year. All LD vehicles.	74
Figure 5.1	Inspection and Maintenance Program Evaluation Methodology.	79
Figure 5.2	CO Reduction in Percentiles. Observed vs. Predicted by MOVES. All Light Duty Vehicles.	80
Figure 5.3	NOx Reduction in Percentiles. Observed vs. Predicted by MOVES. All Light Duty Vehicles.	80
Figure A.1	Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.	107
Figure A.2	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.	110

Figure A.3	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. LDT2 Trucks Only.	110
Figure A.4	Seasonal Average Values of NOx. All Passenger Vehicles (LDT & LDV). Atlanta 13 Counties.	111
Figure B.1	Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.	142
Figure B.2	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.	145
Figure B.3	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. LDT2 Trucks Only.	145
Figure B.4	Seasonal Average Values of NOx. Passenger Cars Only.	146
Figure C.1	Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.	178
Figure C.2	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.	181
Figure C.3	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. LDT2 Trucks Only.	181
Figure C.4	Seasonal Average Values of NOx. Passenger Cars Only.	182
Figure D.1	Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.	213
Figure D.2	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.	216
Figure D.3	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Trucks Only.	217
Figure D.4	Seasonal Average Values of NOx. Passenger Cars Only.	218
Figure E.1	Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.	248
Figure E.2	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.	251
Figure E.3	Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Trucks Only.	251
Figure E.4	Seasonal Average Values of NOx. Passenger Cars Only.	252
Figure E.5	Average CO Levels. Atlanta 13 Counties Area and Augusta-Macon. Passenger Vehicles Only.	254

LIST OF SYMBOLS AND ABBREVIATIONS

ANOVA	Analysis of Variance
AQG	Air Quality Group
ASM	Acceleration Simulation Mode
CAA	Clean Air Act
CAAA	Clean Air Act Amendments
CAFÉ	Continuous Atlanta Fleet Evaluation
CO	Carbon Monoxide
CO2	Carbon Dioxide
GRD	Georgia Registration database
GTRI	Georgia Tech Research Institute
GVWR	Gross Vehicle Weight
HC	Hydrocarbons
I/M	Inspection and Maintenance
LDT	Light Duty Truck
LDV	Light Duty Vehicle
Lower CL Dif	Shows the lower confidence limit for the difference
MOBILE6.2	Vehicle Emission Modeling Software
MOVES	Motor Vehicle Emission Simulator
NOx	Oxides of Nitrogen
OBD	On Board Diagnostics
OBD II	On Board Diagnostics current generation
PPM	Parts Per Million

Prob>|t|	The p-value for the parameter.
RSD	Remote Sensing Device
RSD3000	Remote Sensing Device 3000 manufactured by ESP, Inc
RV	Recreational Vehicle
SIP	State Implementation Plan
Std Dev	Standard Deviation
Std Err Dif	Shows the Standard Error of the Difference
t Ratio	The t Statistic for the parameter, computed as Estimate/Std Error
US EPA	United State Environmental Protection Agency
VIN	Vehicle Identification Number
VMT	Vehicle Miles Traveled
VOC	Volatile Organic Compounds
VSP	Vehicle Specific Power

SUMMARY

The reference method for evaluating vehicle inspection/maintenance programs yields several advantages over other methods using on road remote sensing data. The reference method could be repeated over time to measure incremental effectiveness as more of the fleet is tested, inspectors become adept at identifying noncompliant vehicles, repair technicians gain experience at repairing emission control failures, and (more pessimistically) motorists learn better how to co-opt the test.

This study addresses the issues associated with applicability of the reference method to the Atlanta I/M evaluations. Currently the method is employed under assumption that experimental and reference fleets behave similarly if normalized by model year distribution and VMT. However, it is not the case. In addition to age and mileage accumulation, emissions from vehicle fleets can be influenced by differences in a variety of other factors including the differences in driving patterns and socioeconomic factors that can impact fleet composition and maintenance trends.

Determination and assessment of factors responsible for emission fluctuations in vehicles of the same type and age group is a primary objective of this research. This is a challenging task since discrepancies in maintenance trends, socioeconomic conditions and vehicle quality are difficult to control for.

A series of “null” experiments were conducted to evaluate factors that might affect emissions across vehicle fleets of the same model year distribution. The emission differences among Atlanta 13 counties fleets were addressed and examined to rule out arguments related to different fuel supply, presence of I/M testing and mileage accumulation. This study examined the effect of different driving conditions on emission levels through the concept of vehicle - specific power (VSP) and confirmed that vehicles driven within same VSP range do not have statistically significant differences in emission rates. Socioeconomic factors, as it turns out, have a lesser effect on emission levels, and can be mitigated by normalizing the model year distributions. As a result, a new correction factor was added to evaluation of Atlanta 13 counties I/M program.

CHAPTER 1. INTRODUCTION

Shortly after passage of the 1990 Clean Air Act Amendments, the US Environmental Protection Agency (EPA) issued regulatory governing motor vehicle inspection and maintenance (I/M) programs. Driven by persistent growth in vehicle travel and chronic air pollution, many of the nation's largest metropolitan areas employed advanced testing technologies and procedures as a way to ensure the efficient operation of vehicle emission control systems.

The inspection process, which applies to light duty vehicles (LDV) of certain ages operated by gasoline engines, involves scheduled or unscheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls. Over the years, I/M testing techniques have undergone significant changes as new technologies have kept being developed and implemented by car manufacturers. In the past, the tests were performed annually or biennially, while a vehicle was idling or was placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of on road driving conditions. Currently, such tests are performed only on vehicles that had been built prior to 1996. Tests of the later vehicles models equipped with advanced onboard diagnostic (OBD II) systems usually involve checking on the reported status of their OBD systems data.

In order to ensure smooth and effective I/M program implementation, current law requires biennial evaluation of enhanced I/M programs and on road measurement of inspected fleet emissions (CAA Title I §182c3C; CAA Title I §182c3Bi; Title I§182c3Ci).

The National Research Council (NRC) has recommended that I/M programs need to be evaluated using on road emissions data collected by remote sensing devices (RSD) (National Research Council, 2001). RSD uses infrared and ultraviolet technology to measure the emissions from the in-use vehicles.[1] The NRC report cited several advantages of RSD data for I/M evaluation. First, the RSD based approach is a cost-effective source of evaluation data compared with the higher per-vehicle costs of the advanced dynamometer testing - the original evaluation approach recommended by federal regulators and conducted on a small sample of vehicles. The RSD data is capable of capturing trends that cannot be discerned through internal inspection records alone, such as motorists avoiding the program and pre-inspection maintenance behavior. Finally, in addition to I/M evaluation, the RSD data can also be used for a variety of purposes, including mobile source emission inventories, "clean-screen" programs that exempt low-emission vehicles from subsequent I/M testing, and high-emitter programs that target polluting vehicles for off-cycle inspection and repair.

In July 2001, the U.S. Environmental Protection Agency (EPA) released a draft guidelines for the use of remote sensing data for I/M program evaluation (U.S. Environmental Protection Agency, 2001). The document outlines equipment specifications and measurement procedures along with the study design techniques and quality control

[1] Infrared technology is used to measure carbon monoxide and volatile organic compounds. Ultraviolet technology is used to measure nitrogen oxides

measures. In addition, the EPA document also discusses three methodologies for analyzing the remote sensing data to determine the I/M program effectiveness which are the “comprehensive, step, and reference methods. The *comprehensive method* compares the on road emissions of the vehicle fleet before and after scheduled I/M testing. The *step method* compares the inspected model year emissions with the uninspected model year emissions during the first year of a new or upgraded I/M program,. The *reference method* compares the emissions of the vehicle fleet located in a particular I/M area with that of a distantly located non-I/M area.

This work is focused on the issues associated with application of the reference method to I/M program evaluation with the assessment of Atlanta, GA I/M program is performed as an illustrative example of such evaluation. The Atlanta metropolitan area is one of the major metropolitan areas in the southeastern United States and is home to thirteen counties with “serious” nonattainment of the federal ozone standard.[2] The Atlanta enhanced I/M program was implemented in October 1996 in this thirteen-county area, replacing a basic I/M program that had been operating in four of the thirteen counties since the early 1980s. The effectiveness of the new I/M program is estimated by comparing the RSD emissions of a sample of its inspected vehicles with that of a sample of vehicles registered in the Georgia cities of Augusta and Macon that are not subject to vehicle testing.

[2] In 2015 the federal ozone standard was proposed to be changed from 0.075 ppm to 0.07 ppm averaged over an eight-hour basis over 3 years. Final rule regarding eight-hour requirement was signed on November 7th, 2018. Atlanta is still required to meet this standard.

The latter areas have demographics, climate and fleet characteristics similar to Atlanta, but do not operate I/M program. The emissions difference in the inspected Atlanta and uninspected Augusta/Macon vehicle fleets are then compared with that predicted by EPA approved emissions models, currently the *Motor Vehicle Emissions Simulator* (MOVES). Viewing model-predicted emissions differences in inspected and uninspected vehicles as the Atlanta I/M program goal and observed on road emission differences as actual program performance, I/M effectiveness is estimated as the ratio of these two estimates.

The next section provides background on I/M programs, including an overview of I/M program operations, and a history of I/M programs in Atlanta and reviews current enhanced I/M evaluation approaches (including the RSD methods outlined in EPA guidance) and their respective strengths and weaknesses. The following chapters describe reference method data sources, methodology and analysis plan, examine research results and discuss conclusions and contributions from this research.

CHAPTER 2. BACKGROUND

2.1 Vehicle Inspection/Maintenance Programs

Vehicle inspection/maintenance (I/M) programs seek, first and foremost, to ensure the effectiveness of vehicle emission control systems. The inspection process involves scheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls.[3] Inspections can be provided by decentralized test-and-repair networks, which allow service stations and automotive repair shops to perform emissions tests and repair failed vehicles, or by centralized test-only networks, in which a limited number of centrally operated facilities perform testing as the sole service. Depending on program design, the test may be performed annually or biennially, while the vehicle idles or is placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of on road driving conditions or, more commonly query the onboard diagnostic system for conditions of the emissions control system. Most programs also have provisions for evaluating the conditions of the evaporative emissions control system, usually through testing the pressure-integrity of the vehicles fuel-tank (gas) cap.

In typical programs, motorists must repair failed vehicles, comprising the maintenance component of the program. Vehicles with repair costs above a set amount

[3] The model year vehicles subject to testing vary across I/M programs.

may qualify for a waiver -- an exemption from further repair and testing -- provided that attempted repairs show some emissions improvement and are not triggered by tampering. Compliance is typically verified through the presence of a vehicle windshield sticker received after passing the test or through the vehicle registration process that requires an emissions certificate.

2.2 Inspection/Maintenance in Atlanta

Atlanta's first I/M program were established in 1981, covering the three ozone nonattainment area counties of Fulton, Cobb and DeKalb. The fast-growing Gwinnett County was added in 1986. The program was implemented through a decentralized test-and-repair network which allowed repair shops, service stations and automobile dealers to perform emission inspections and emissions-related repairs. Testing was originally required for the latest ten model year vehicles, but was expanded in 1986 to include the latest twelve model years. To receive an emissions compliance certificate, cars were required to pass an idle emissions test and an inspection of the catalyst, air pump and fuel inlet restrictor for evidence of tampering. Owners of failing cars that spent more than $50 for repairs qualified for a waiver and an emissions certificate, so long as repairs were not due to tampering and showed some emissions improvement. Owners of cars that failed the tampering inspection were required to obtain repairs to bring their emissions into compliance regardless of cost.

In response to the 1990 Clean Air Act Amendments (CAAA), the Georgia legislature revisited emissions testing in 1992[4]. This legislation enabled the Georgia Department of Natural Resources (GDNR) to upgrade Georgia's I/M program to an "enhanced" program, bringing it into compliance with the 1990 CAAA and new federal I/M federal regulations. This enhanced version of the program received limited implementation in October 1996[5], with emission inspections required only for those vehicles migrating to the Atlanta I/M program area. The new program commenced in January 1997, with biennial emissions testing required of all vehicles from the 1975 model year up until two years of age. The new program also spanned the 13-county nonattainment area, incorporating nine new counties that were not subject to the previous basic I/M program.

After the initial two years of operation, several changes were made to the program. Notably, vehicles over six years of age at the time had to undergo the more rigorous Acceleration Simulation Mode (ASM) testing beginning in October 1998 (In the initial period all vehicles were subjected to a two-speed idle test (TSI)[6]). Secondly, vehicles that failed emissions testing were required to be brought into compliance by repair. Owners of covered vehicles in the 13-county ozone nonattainment area were now required to show

[4] 1992 Georgia Air Quality Act, Article 2: Motor Vehicle Emissions Inspection and Maintenance Act (OCGA Section 12-9-40 et seq.).

[5] October 1996 was chosen as the soonest possible start-up date after the previous basic I/M program, which operated during a January-to-April vehicle registration "season." Vehicle registration is now conducted year-round in Georgia, as is enhanced emissions testing.

[6] Two-speed idle (TSI) testing procedure that measures emissions under idle and a 2500 RPM engine speed

proof of a passing emissions inspection, a waiver, or proof that they qualify for an exemption in order to register their vehicle.

The proposed research concentrates on the sixth and seventh two-year periods of inspection and maintenance program operation in Atlanta and covers years from 2007 through 2010. By this time certain significant changes have been made to the Atlanta enhanced I/M program. The waiver limit[7] had been increased several times. In 2001, testing frequency changed from biennial to annual; the requirement to inspect vehicles back to 1975 model years was replaced with the requirement to inspect the latest 25 model years; and the exemption from testing of the newest two model years was changed to exemption of the newest three model years. ASM tests are performed only on vehicles that had been built prior to 1996. Test of later vehicles equipped with advanced onboard diagnostic (OBD II) systems were tested by evaluation of their OBD records.

2.3 Existing Enhanced I/M Program Evaluation Methods

Three types of data currently dominate the evaluation of enhanced I/M programs: I/M records, which document the results of each inspection; roadside pullovers, which administer emissions tests to vehicles of randomly selected willing motorists; and remote

[7] In 2019 the Repair Waiver limit amount is set on $918, Georgia's Clean Air Force, www.cleanairforce.com

sensing data, which measures on road vehicle emissions. This section reviews evaluations employing each data type, along with the strengths and weakness of each.

2.3.1 Evaluation based on Emission Inspection Records:

Emission inspection records are the sole basis for the records-type I/M evaluation. I/M test records provide a cost-effective source of evaluation data because they are routinely generated and easily accessible. Because I/M records cover the entire inspected vehicle population, statistical conclusion validity is generally not an issue: evaluators can control for a variety of vehicle characteristics that influence emissions. The availability of odometer data in most I/M records is also advantageous, enabling evaluators to control for the influence of mileage on emissions. A final advantage stems from inspection/maintenance protocols, which are designed to correlate with the Federal Test Procedure[8] and to facilitate quality control.

However, I/M records suffer from weaknesses that limit their reliability as the sole indicator of program performance. One of them is the inability to parcel out fraudulent testing behavior, particularly when inspectors substitute clean-emitting vehicles for unrepaired high-emitting ones on the retest (Wenzel, Singer, & Slott, Some Issues in the Statistical Analysis of Vehicle Emissions, 2000). I/M records may also underestimate

[8] The Federal Test Procedure is an elaborate testing protocol established in the early 1970s to certify manufacturer compliance with the 1970 Clean Air Act-mandated new vehicle emission standards.

program effectiveness by missing pre-inspection maintenance performed by some motorists to lower I/M test failure risks. While it is difficult to quantify the impact of such maintenance, it is expected to yield artificially low baseline emissions and thus underestimate I/M program effectiveness. Generally speaking, these weaknesses speak to the role of I/M records as an internal, not an independent, source of evaluation data.

Evaluations employing I/M records also make tradeoffs between internal validity and representativeness of the data. The inspection process employs highly-controlled conditions to ensure that vehicles are measured under consistent circumstances (e.g., engine stress, vehicle speed, and temperature). While these controls reduce confounding influences on emissions, they represent only a fraction of driving conditions that typify onroad driving. Consequently, the ability to extrapolate I/M test emissions to onroad emissions is limited.

To estimate I/M effectiveness, some evaluations calculate the average emissions difference between the initial and final test scores on failing vehicles and assume that the difference is attributable to the I/M program. Three studies used this approach to evaluate different time periods of the Arizona enhanced I/M program. Two of these studies (Wenzel, 1999; Glover & Brzezinski, 1997) estimated a 14 percent reduction in carbon monoxide (CO), a 15 percent reduction in hydrocarbons (HC), and a seven percent reduction in nitrogen oxides (NOx). The third study (Ando, McConnell, & Harrington, 1999), focusing on repaired vehicles, estimated emission reductions of eight, eight and fourteen percent for CO, HC and NOx.

Sierra Research (Sierra Research, 1998) also compared initial and final emission results for failed vehicles in AirCare, the Canadian Vancouver/British Columbia emissions testing program. This study estimated I/M emission reductions of 13 percent for CO, 9 percent for HC, and 4 percent for NOx. Replacing the initial emission results of failed inspections with EPA model predictions of an untested fleet's emissions, the researchers estimated 16 percent, 20 percent, and 14 percent emission reductions for CO, HC and NOx respectively. These latter emission differences are thought to be higher than the former because model predictions, as opposed to initial inspection results, are not influenced by pre-inspection maintenance behavior.

The Colorado enhanced I/M program was twice evaluated using inspection records. The first analysis, comparing final test scores for vehicles inspected in 1997 with the new program's first 2,138 initial inspection test scores in 1995, indicated CO emission reductions in the range of 30 to 34 percent (Environ, Inc., 2003). The second analysis compared failed vehicles' initial and final inspection results from 1998 that had been converted to Federal Test Procedure scores. The comparison, which normalized repair benefits to the entire inspected fleet, suggested that CO had been reduced by eight percent and HC by six percent, with NOx increasing by one percent (Environ, Inc., 2003) While the study results seem contradictory, they cover different timeframes, make divergent assumptions (about deterioration rates, the fate of vehicles with final failures) and employ different measures in estimating I/M effectiveness.

One weakness in attributing before-after emission differences to I/M is the potential for "regression to the mean" emissions behavior, in which a portion of I/M failures will

register lower emissions on the final inspection without repair.[9] This phenomenon is driven by tremendous emissions test-to-test variability, the presence of vehicles with marginally failing emissions, and variance in environmental conditions favorable to test performance. Without verifying repairs, the emissions differences between initial and final test scores may overestimate program effectiveness.

2.3.2 Evaluation based on Roadside Emission Inspections:

Used exclusively in California, roadside emissions tests are administered with the aid of law enforcement officers who randomly pull vehicles over and ask motorists to voluntarily submit their vehicles to an emissions inspection. Volunteer license plate numbers are then used to query the I/M program database to determine those vehicles with and without an inspection in the past twelve months. Recently inspected and uninspected vehicle emissions are then compared to estimate the emission reductions due to enhanced I/M. Roadside emissions estimates of 1999 enhanced I/M program effectiveness indicate emission reductions of 13 percent for CO, 14 percent for HC, and 6 percent for NO (California Air Resources Board, 2000).

[9] Regression to the mean occurs when two imperfectly correlated measures are compared for a nonrandom sample. The nonrandom sample is typically drawn from high or low scorers on either measure. Regression to the mean occurs when the sample mean moves towards the population mean in the absence of intervention. In the context of I/M evaluation, this means that certain vehicles failing their initial I/M test will score more closely to the mean of the population on the retest, i.e., register passing emissions, without repair. Regression-to-the mean can also occur in vehicles that pass their initial inspection but would fail a subsequent retest.

In 2009, the California Air Resources Board, in cooperation with Bureau of Automotive Repair, hired Sierra Research Inc. to conduct an independent research and analysis of the Smog Check Program using data collected from roadside inspections conducted in 2003-2006. The study compares roadside inspection results for 1976-95 (pre-OBDII) model year vehicles to the Smog Check inspection results reported by Smog Check stations. The study concluded that of the 1976-95 vehicles sampled, 19% of the vehicles initially passed a tailpipe inspection at a licensed Smog Check station, but failed a roadside audit inspection within a year. The data also showed that 49% of the vehicles that failed a roadside audit inspection had failed and then subsequently passed emission test at a Smog check station (California Air Resource Board and Bureau of Automotive Repair, 2010).

As with I/M program data, roadside pullovers enable the collection of odometer data for mileage estimates. In contrast with I/M program data, the spontaneity of roadside inspections preclude fraudulent test results that overestimate effectiveness, as well as pre-inspection maintenance behavior that underestimates program effectiveness. However, because roadside emissions tests employ a portable version of official inspection procedures, they sacrifice real-world driving conditions. Furthermore, the approach is costly and generates limited data, requiring as many as four technicians and one law enforcement officer to measure approximately 25 vehicles per day (Wenzel, Gumerman, Singer, & Sawyer, 2000). Self-selection bias is a risk because the test is voluntary and tends to yield a ten percent refusal rate (Wenzel, Gumerman, Singer, & Sawyer, 2000).[10]

2.3.3 Evaluation Methods based on Data from Onroad Vehicles:

Data collected by remote sensing devices (RSD) that measure the emissions of vehicles while they are being driven. The advantage of in-transit measurement is the ability to observe a vehicle's emissions under typical driving conditions, which cannot be as easily captured by traditional controlled emissions testing procedures. Remote sensors can measure a large number of vehicles, an important attribute given the need to control for tremendous emissions variability due to vehicle type, age, make and model, and emission control technology. A final advantage stems from the unscheduled nature of the measurement, which precludes pre-inspection and fraudulent maintenance behavior that can occur when motorists (as with I/M tests) know when a measurement will occur.

In contrast with the highly controlled parameters of the emissions inspection, the physical circumstances of remote sensing data collection are only approximated through sampling site characteristics (e.g., moderate grades to ensure vehicles operate under only a slight engine load and sampling sites that avoid residential areas to minimize inflated emissions from cold engines). Another drawback is that remote sensors capture a split-

[10] The evidence of such bias is mixed. One recent study that used remote sensing to measure the vehicle emissions of refusals and participants alike found no significant difference between the two groups (Wenzel et al, 2000, pg. III-8), while an earlier similar study found that refusal vehicles had 2.5 times the emissions of volunteer vehicles (Stedman, 1994).

second emissions reading that may not reflect a vehicle's typical emissions, making larger samples sizes a requirement to average out random emission fluctuations and to profile emissions aggregated within vehicle type (cars vs. trucks) and model year.

Remote sensing data has been used in three ways to evaluate I/M programs.

2.3.3.1 Comprehensive Method

The first method averages the emissions of vehicles measured before initial and after final I/M testing, with the difference attributed to I/M program effectiveness. Dubbed the "comprehensive method" in EPA evaluation guidance, emissions differences can also be generated for various subfleets, such as vehicles initially failing and ultimately passing I/M testing versus failing vehicles that never receiving a final pass. This approach enables a variety of I/M-related analyses, such as deterioration rates of I/M repairs, the influence of pre-I/M repairs on emissions baselines, and a comparison with estimates based on I/M records alone. The major disadvantage to this approach is the enormous volume of onroad data required to measure a representative sample of vehicles before and after I/M testing. Sample size requirements hinge on the probability of measuring vehicles onroad within a specific time period of I/M testing, a probability that fluctuates with testing frequency and the distribution of sampling throughout the year.

The comprehensive method was used to estimate the effectiveness of the California South Coast Air Basin's enhanced I/M program in 1999 (Wenzel, Gumerman, Singer, & Sawyer, 2000). "Smog Check" I/M records were used to delineate tested from untested vehicles by the existence of an enhanced inspection within the past twelve months.[11] A comparison of these vehicle groups indicates a ten percent reduction in CO, a four percent reduction in HC, and a five percent increase in NOx. An earlier remote sensing study in California in 1996 compared the onroad emissions of 3.5 million vehicles 30 to 90 days before with up to 90 days after their basic I/M test (Klausmeier & Weyn., 1997). For those vehicles that failed their initial smog check and then passed, both CO and HC emission differences registered at 20 percent. Normalizing this result to the entire fleet yielded an estimated nine percent emissions reduction in HC and CO. A third evaluation, of the Arizona enhanced I/M program in 1997, analyzed four million remote sensing measurements on 1.2 million vehicles in the Phoenix I/M area (Wenzel, 1999). The results indicated a seven percent reduction in CO and an 11 percent reduction in HC.

One weakness in the comprehensive method is the potential seasonal effects that results from the year-round testing required to obtain adequately sized samples. Users of this method have also tended to rely on a few high-volume sites, yielding a large number of repeat vehicles that lower the fraction of unique vehicles that could be reached at a greater number of sites.

[11] Untested vehicles may have been inspected under the previous basic I/M program more than twelve months ago or they may have had an enhanced inspection after the remote sensing reading.

2.3.3.2 Step Method

A second I/M evaluation approach using remote sensing, known as the Step Method, compares inspected with uninspected vehicles during the first year of a new or upgraded program. The uninspected vehicles comprise an internal control group against which to compare the emission reductions of the inspected vehicles. Because the method applies to the early phases of a new or improved program, it can be used only once to assess program effectiveness.

A remote sensing study of the Colorado Enhanced I/M program compared odd (inspected) and even (uninspected) model year vehicles during the end of the first year of a new biennial enhanced I/M program (Stedman D. , Bishop, Aldrete, & Slott, 1997). At that point, in program history, all odd model year vehicles should have been inspected, whereas all even model year vehicles had no reason to be inspected. This timing rendered even model year vehicles the untested control group against which to compare the odd model year vehicle emissions. The comparison of odd and even model year emissions suggested that Colorado's enhanced I/M program had reduced CO between five and nine percent, while HC and NO showed no improvement.

Three factors limit the generalizability of the Colorado study results to enhanced I/M program effectiveness. Remote sensing took place in a single location, which avoids any confounding socioeconomic or physical influences at different sites but limits generalizability to the overall fleet. Furthermore, vehicles traveling past the remote sensing site were decelerating, which does not represent typical driving conditions and is not the optimal condition for measuring carbon monoxide (Environ, Inc., 2003), p. 2-19). A third limitation was that the study measured vehicles transitioning from an annual basic I/M program to an enhanced I/M program, rendering it an evaluation of incremental program effectiveness and not a complete estimate of enhanced I/M program performance.

The research that replicates the original Denver Step Method analysis was made by Air Quality Laboratory at Georgia Institute of Technology for a 1997 Atlanta I/M program. This evaluation was conducted separately for the nine outlying Atlanta counties and the four counties that are closest to the center of the city. The results of the analysis are similar to those found by Stedman et al. in Denver. While the Denver carbon monoxide (CO) weighted program benefit was 6.9%, the Atlanta area CO weighted program benefit is found to be 11.5% and 4.9% for the nine-county and four-county Atlanta areas, respectively. The study concluded that the 1997 I/M program change in Atlanta yielded a noteworthy and observable change in fleet emissions (Corley, DeHart-Davis, Lindner, & Rodgers, 2003).

2.3.3.3 Reference Method

A third approach using remote sensing data (the one used in this study) compares the onroad emissions of vehicles registered in an I/M area to that of vehicles registered in non-I/M areas. The non-I/M area serves as a surrogate untested fleet. The validity of this approach relies on the selection of a non-I/M area comparable in fleet age distribution, a well-documented contributor to vehicle emissions; climate, which can accelerate emission control equipment deterioration; and demographics, which influences the age, quality, and maintenance of the vehicle fleet. This approach was originally applied to the basic I/M program operating in four counties of the thirteen-county Atlanta ozone nonattainment area, with the nine nonattainment counties without I/M comprising the untested fleet. The analysis indicates that car and truck emissions for CO were fifteen and ten percent higher, respectively, in the uninspected nine-county fleet than in the inspected four-county basic I/M fleet. The study is limited by its inability to control for differences in mileage and socioeconomic conditions between the two vehicle fleets.

The same method was also used to evaluate Atlanta Biennial I/M program by comparing Atlanta 13 counties inspected fleet with Augusta-Macon untested fleet. Assuming that on-road emissions differences represent observed effectiveness and EPA approved model-predicted emissions differences represent effectiveness goals, this study had shown that the Atlanta enhanced I/M program appears to be achieving 83% of its targeted emissions reductions (DeHart-Davis, Corley, & Rodgers, 2002). Later on that approach became a basis for further biennial evaluations of Atlanta enhanced I/M programs (see Appendices).

CHAPTER 3. RESEARCH DESIGN COMPONENTS

3.1 Description of Research Setting

This research was conducted in the metro area of Atlanta and two other Georgia cities: Augusta and Macon, that are similar by socioeconomic and travel pattern. A supplementary statute of Georgia's State Implementation Plan (SIP) for the Atlanta Ozone Non-attainment Area details the Atlanta's I/M program. This plan was based on Atlanta's classification as a "serious" non-attainment area and thus an Enhanced I/M program was therefore a mandatory requirement for metropolitan Atlanta under provisions of the Clean Air Act Amendments (CAAA) of 1990.

Originally, EPA specified that ozone non-attainment areas classified as "serious" or above would be automatically applied to the entire Metropolitan Statistical Area, "unless the state could demonstrate that such action would not be appropriate". The Atlanta Metropolitan Statistical Area currently consists of 21 counties. However, only 13 of these were decided to be the subject of I/M program. Those counties had been chosen on the basis of the following three criteria (SIP,1999):

1. Population density, urbanization, commuting patterns, population increases, etc.
2. The emission density of stationary sources and the density of mobile sources expressed as VMT (Vehicles Miles Traveled).
3. Meteorological factors, biogenic (natural) vs. anthropogenic (man-made) emissions and physical boundaries that may influence movements of precursor pollutants.

3.2 On Road Emission Data Collection

Data for this research were collected under the Continuous Atlanta Fleet Evaluation (CAFE) program. These data provide on road emissions readings that are used to evaluate Atlanta enhanced I/M program performance. CAFE uses remote sensing devices to annually measure the emissions of in-use vehicles in the 13-county I/M program area, and periodically in two cities located more than 75 miles from Atlanta that do not require vehicle emissions testing.[12] The study is an ongoing effort started in 1993 to collect vehicle emissions data for assessing a variety of trends, including fleet turnover, emission control deterioration, and socioeconomic impacts of mobile source control strategies.

RSD measures the emissions of passing vehicles remotely and unobtrusively so motorists are minimally aware of the equipment and do not alter their natural driving behavior. To that end, the remote sensing instrumentation is housed in a van parked on the roadside along with a video camera. An infrared light source and its generator are placed on the opposite side of the road or on the median to create a beam of light that traverses the road. When a passing vehicle breaks the beam, it triggers a measurement of hydrocarbons, carbon monoxide, and nitrogen oxides in the exhaust. Simultaneously, a video camera records the vehicle's license plate, which is automatically scanned into the database of emissions measurements.

[12] Augusta is located 136 miles east of Atlanta, whereas Macon is 76 miles south of Atlanta.

Remote sensing works the best if the subject vehicles are operating in a predictable manner and the vehicles' engines are operating under a moderate continuous load. This is generally achieved through the careful selection of sites at which remote sensing equipment is being installed. Site selection is a crucial aspect of any study that employs remote-sensing methods.

Remote sensing sites for this study were selected based on following criteria:

1. Safety: the site should be safe for both drivers and operators (e.g. adequate site distance and safe access) and have sufficient space on both sites of the roadway to safely place the equipment.
2. The site should satisfy certain geometric criteria: the sites should have road geometries and vehicle operating modes compatible with and desirable for remote sensing (e.g. single lane operation, moderate vehicle specific powers, and absence of cold start emissions). This is necessary to guarantee that the vehicles are operating at moderate speeds either on a small positive grade and/or with modest positive accelerations. To achieve continuous load on the engine, small positive grades are generally more desirable. These conditions create observed vehicle specific powers (VSP) in the desirable range (5-25 kW/Tonne) for remote sensing measurements.
3. Sites should be geographically located in areas that are desirable for a particular study in terms of demographics and fleet composition.

These site selection criteria and the sites selected for the Atlanta CAFÉ program are discussed extensively by Samoylov (Samoylov, 2013).

Figure 3.1 illustrates the distribution of measurement sites for the 2010 measurement year of the Atlanta CAFÉ program.

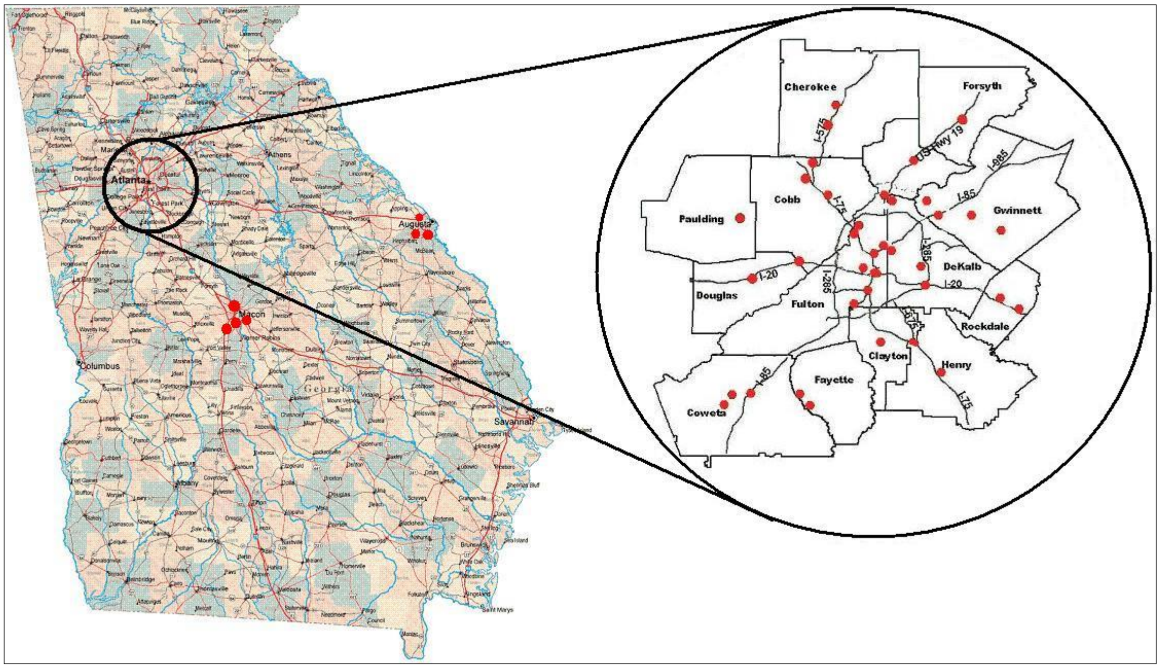

Figure 3.1 2009- 2010 Remote sensing sites for the Continuous Atlanta Fleet Evaluation (CAFÉ) in the thirteen-county Atlanta vehicle inspection and maintenance (I/M) area and Augusta and Macon, Georgia

After data collection, remote sensing measurements are merged with vehicle registration records using the vehicle license plate. The resulting database allows various characteristics of measured vehicles to be identified, including vehicle identification

number,[13] make, model year, and vehicle type. License plates are also linked with inspection/maintenance records to identify vehicles with prior emission inspections.

RSD sampling sites are selected to ensure physically consistent but demographically diverse characteristics. Single straight lines of traffic with an average 35 mile-per-hour velocity are sought to facilitate single vehicle measurements and speeds that maximize measurement opportunities. Driver behavior and driving maneuvers are also observed at each site to ensure that remote sensing measurements would not be biased high by acceleration or low by coasting. Finally, notations are made during the site visits regarding any obvious or suspected diurnal patterns that exist which affect the traffic volume. If distinct variations are found to exist in sites ultimately selected, sampling times are scheduled to account for those diurnal patterns. U.S. Census tract data and traffic count reports inform the selection of different income ranges and land uses.

The remote sensing sites relevant to this study reside within the 13-county Atlanta I/M program area, 12 Atlanta counties without an I/M program but subject to the Atlanta clean fuel program, as well as the Georgia cities of Augusta and Macon that have neither program. The latter locales do not require emissions testing and thus provide an

[13] Vehicle identification numbers are 17-digit alphanumeric strings that uniquely identify every vehicle manufactured. When decoded, they provide additional characteristics on vehicles. The VIN-decoded data of particular relevance to this research are vehicle type (car, truck, multi-purpose vehicle, van) and model year.

uninspected vehicle fleet to serve as a control group. These cities were chosen after a review of census data and registration records revealed them to have characteristics – median household income, population density, and fleet distribution -- most similar to Atlanta than three other Georgia cities considered. But for the reasons which will be explained later for the present analysis we also used as the reference point the data collected from the vehicles registered in 12 counties that surround Atlanta I/M program area.

According to the state regulation, effective April 1, 1999 the sulfur content of all gasoline supplied in a 25 Atlanta region[14] shall not exceed a seasonal average of 150 ppm (by weight) and, effective April 1, 2001, a per-gallon cap of 500 ppm (by weight)[15]. This rule made vehicle operational conditions in Atlanta 13-county nonattainment area and Augusta-Macon significantly unequal. Subsequent changes in federal regulations lowered this cap to 10 ppm. Since there is no mechanism to separate benefits received from the usage of low sulfur gasoline and emission reductions due to I/M program, the usage of Augusta-Macon fleet as a control group for I/M program evaluation became questionable while there were differences in the emissions characteristics of the fuels between the regions (since resolved by changes to the federal program). In an effort to eliminate the fuel effect during these intervening years, the data collected from the vehicles registered in

[14] 25-county Atlanta region include 13-county I/M program area and 12 additional counties without I/M program: Barrow, Bartow, Butts, Carroll, Dawson, Hall, Haralson, Jackson, Newton, Pickens, Spalding, Walton.

[15] Rules for Air Quality Control Chapter 391-3-1, July 20, 2005 http://www.gaepd.org/Files_PDF/rules/rules_exist/391-3-1.pdf

twelve counties[16] that are not subject to the I/M program but receive the same fuel as Atlanta 13-counties I/M program area have been used in the present analysis as a reference point. The data collected on Augusta-Macon sites represents combined benefits from both I/M and GA fuel programs.

3.3 Research Objective and Justification. General Reference Method. Case of Atlanta I/M Evaluation

The reference method is essentially a comparison of measurements made at experimental and reference conditions. It is assumed that both settings have the same general characteristics so any differences observed could be attributed to the treatment effect. Applied to Atlanta's I/M evaluation this method compares observed emission differences of inspected and uninspected vehicles assuming that both light duty fleets have the same main characteristics (such as model year distribution, etc.). Observed emission differences in its turn are being compared with those predicted by an EPA emissions model (typically MOVES). Equation 1 represents the formula for estimation of I/M Effectiveness is as follows:

---

[16] Eight of these twelve counties are in the Atlanta Metropolitan Statistical Area and thus are considered "Atlanta-area" counties.

$$\text{Effectiveness} = \frac{\sum_{ij}\left[\left(O_{n_{ij}} - O_{m_{ij}}\right)/O_{n_{ij}}\right]\left(P_{n_{ij}}\right)\left(C_{ij}\right)\left(VMT_{ij}\right)}{\sum_{ij}\left(P_{n_{ij}} - P_{m_{ij}}\right)\left(C_{ij}\right)\left(VMT_{ij}\right)} \tag{1}$$

where: O_m and O_n are the average onroad emissions observed for a particular model year and vehicle type for I/M program and non-program vehicles, respectively; P_m and P_n are the model-estimated emission factors for I/M program and non-program vehicles for a given model year a vehicle type; C_{ij} is the fraction of the Atlanta fleet of that model year and vehicle type observed by CAFE; and VMT_{ij} is the average annual vehicle-miles-traveled by model year and vehicle type in the I/M program area.

This formula enables the different units of measurement between on road and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form. Predicted and observed emissions differences in I/M program and non-I/M program vehicles are normalized by model year to the on road fleet fraction and average annual mileage of that model year. This exercise is designed to remove the effect of different vehicle age distribution and miles traveled from the analysis. However, while model year and VMT are definitely important parameters that affect emission levels others important variables remain uncontrolled by current approach. Characteristics that might significantly affect aggregate emission levels are summarized below:

- *Age* – the older is the vehicle, the higher levels of pollutants it emits. As was noted before, that parameter is being controlled by normalization to a common age distribution.
- *Annual Mileage* – controlled by normalization to a common distribution.

- *Income level* – cars being regularly maintained perform better than their counterparts. It is believed that people with higher income maintain their cars better than those with lower incomes. Therefore, significant differences in socioeconomic characteristics between experimental and reference fleet might pose a noticeable bias on I/M evaluation.
- *Ambient Temperature* – ambient temperatures have been reported to affect automotive emissions since 1966 (Spindt, Dizak, Stewart, & Meyer, 1979).
- *Different "Make" distribution* – cars built by certain manufacturers are viewed as less polluting (HONDA for example). If true, fleets with large percentage of cars produced by such manufactures will pollute less emission due to technological advances and not due to maintenance program. Figure 3.2 represents the observed differences in mean CO levels across vehicles manufactured by different companies (it has to be noted, however, that data was not normalized by model year).
- *Driving Conditions* – emissions produced by vehicles operating under different driving conditions (speed, acceleration) vary significantly.

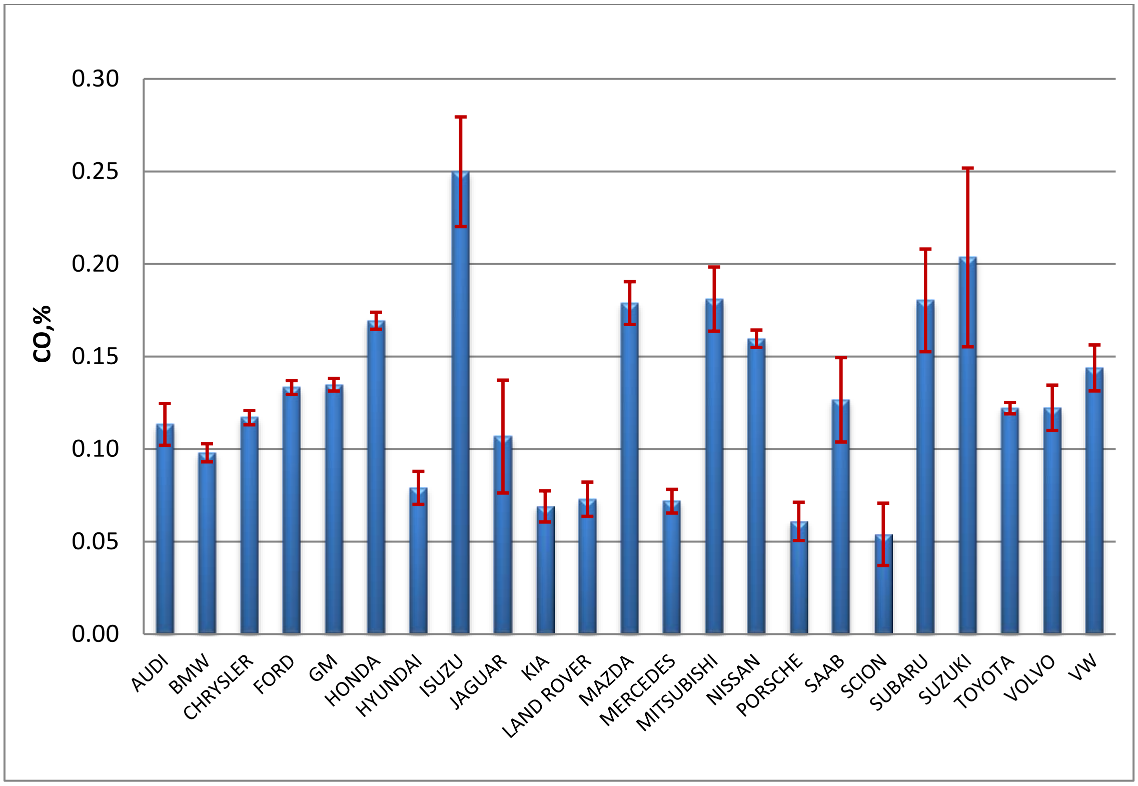

Figure 3.2 Mean CO emission levels for vehicles manufactured by different companies. Measured in Atlanta thirteen-county area in 2010.

Based on the above, we can say that one of the biggest problems associated with application of reference method to the Atlanta I/M evaluation is the construction of a proper control fleet. Determination and assessment of factors responsible for emission fluctuations in vehicles of the same type and age group is a primary objective of this research. Furthermore this study aims to improve current methodology used for Atlanta I/M evaluation and similar reference method studies.

3.4 Research Questions

Following problems and research questions had been investigated during the research presented:

- Which correction factors need to be developed for corresponding fleet to assure its compatibility with experimental fleet?
- To what extent do original vehicle characteristics such as Age, Model and Make affect emission distribution?
- To what extent do different socioeconomic characteristics such as income affect vehicle emissions?
- To what extend driving conditions such as speed and acceleration affect vehicle emissions?

It is hypothesized that:

H1: There are no significant differences in emission levels among vehicles registered in Atlanta 13 Counties.

H2: Vehicles registered in counties with higher median income levels pollute less than those registered in counties with lower median income.

Those hypotheses are to be tested by analyzing the combined database formed from Georgia Registration Database, I/M Data, and RSD data through past years. This research concentrates on data collected in 2010 due to the availability of the greatest quantity of remote sensing readings but also provides results and an overview of average emission levels for other years (see Appendices A-E). Year selection is based on biennial nature of the I/M program evaluation.

3.5 Data Description

For the purpose of this study four data sets were acquired: Inspection and Maintenance Data for the emission tests carried out in 13 county Atlanta area, Georgia Registration Data, Remote Sensing Data and American Community Survey.

Inspection and Maintenance (I/M) Data is available through Department of Natural Resources (DNR). DNR had gathered these databases from the field missions of Clean Air Force (CAF) spread all-over 13 counties of Atlanta. Enhanced I/M program was initiated in Atlanta in October 1996, but the reliable data sets are available starting from 1998. The database for the 2010 measurement year contains about 2 million observations. The following variables from I/M database were used in this study:

1) VIN - vehicle identification number;
2) Test Date - the month/day/year the test was executed;
3) Vehicle Year - the year car was manufactured;
4) Overall Result – the test result (passed/failed/aborted);
5) Vehicle Make – the manufacturer of the vehicle;
6) Vehicle Model – the model of the vehicle.

No hypotheses were made about two last variables, but by monitoring those interesting patterns were observed as was mentioned in “Research Justification” section; for example: between the Make of the vehicle and its emission levels.

Registration databases for years 1998-2010 were acquired from Department of Revenue. The following variables from those records were used:

1) VIN – vehicle identification number;
2) County Code – code of the county in which vehicle is registered;
3) Zip Code – zip code in which vehicle is registered;
4) Vehicle Year – the year car was manufactured;
5) Vehicle Make – the manufacturer of the vehicle;
6) Vehicle Model – the model of the vehicle.

Remote Sensing Data for the research was collected by Air Quality Group (Aerospace, Transportation and Advanced System Laboratory, Georgia Tech Research Institute) under the Continuous Atlanta Fleet Evaluation (CAFE) program. The data collection method and procedures were discussed in previous sections.

Mean Household Income data was acquired through American Community Survey distributed by U.S. Census Bureau.

3.6 Methodology and Analysis Plan

The research design may be classified as an *interrupted time series with multiple replications and nonequivalent no treatment control group*. Presence of the I/M program is considered to be the treatment.

The study consists of the following steps:

First, unique vehicles by VIN should be identified for each annual I/M database. This is important since often one vehicle appears in the same data file several times due to or because it did not pass the test on the first attempt. Multiple appearances in the RSD database are not important since the car receives a unique reading each time. Analyzing the qualitative/dummy variable “overall result”, we identify vehicles, that have undergone the I/M test. We then merge the I/M records with the RSD readings. The resulting database is then partitioned using the “county” variable to produce, for each of the years observed, two groups of cars/trucks:

1. Cars/Trucks registered in Atlanta I/M area form the **experimental group**.
2. Cars/ Trucks registered in Augusta or Macon form the **control group**.

The use of VIN as the key variable in tracking the registration path of vehicles ensures validity and reliability of empirical results since each vehicle normally keeps its Vehicle Identification Number during its whole lifetime.

The comparison of VIN in I/M and Registration databases will help us to determine following variables:

For the experimental group:

X1= number of vehicles that undergone I/M test and registered inside of Atlanta 13 county area during all lengths of observation;

X_i= number of vehicles in other categories we would like to control (County, Make, Model, Income Category, etc.)

For the control groups:

Z1= number of vehicles registered in the control group area during all lengths of observation;

Z_i= number of vehicles in other categories we would like to control (Make, Model, Income Category, etc.)

The RSD database provides pollutant levels (CO, HC, NOx) for all controlled groups of vehicles.

Figure 3.3, Figure 3.4 and Figure 3.5 illustrate steps taken to develop, clean and split data into categories needed for analysis.

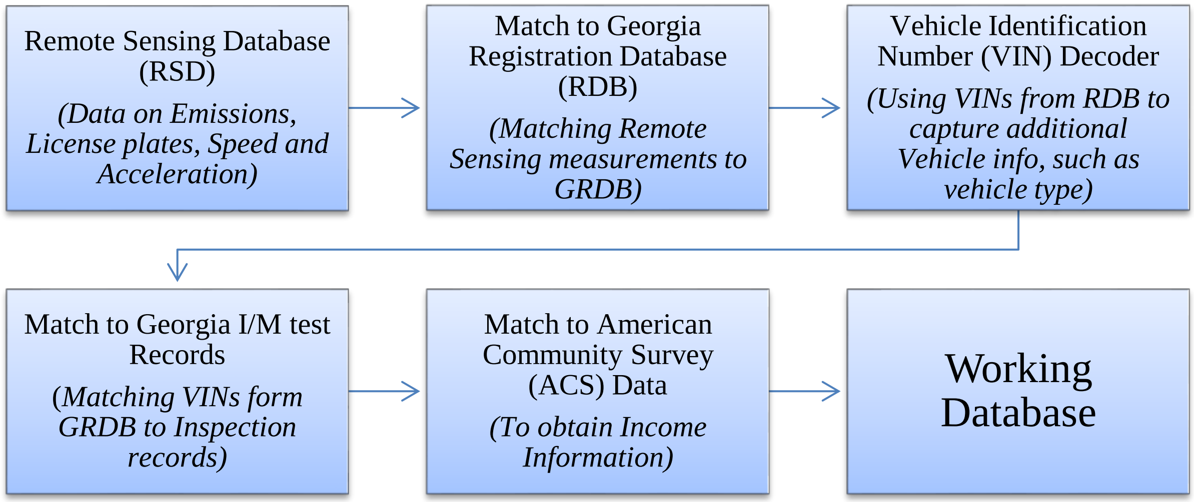

Figure 3.3 Dataset Development.

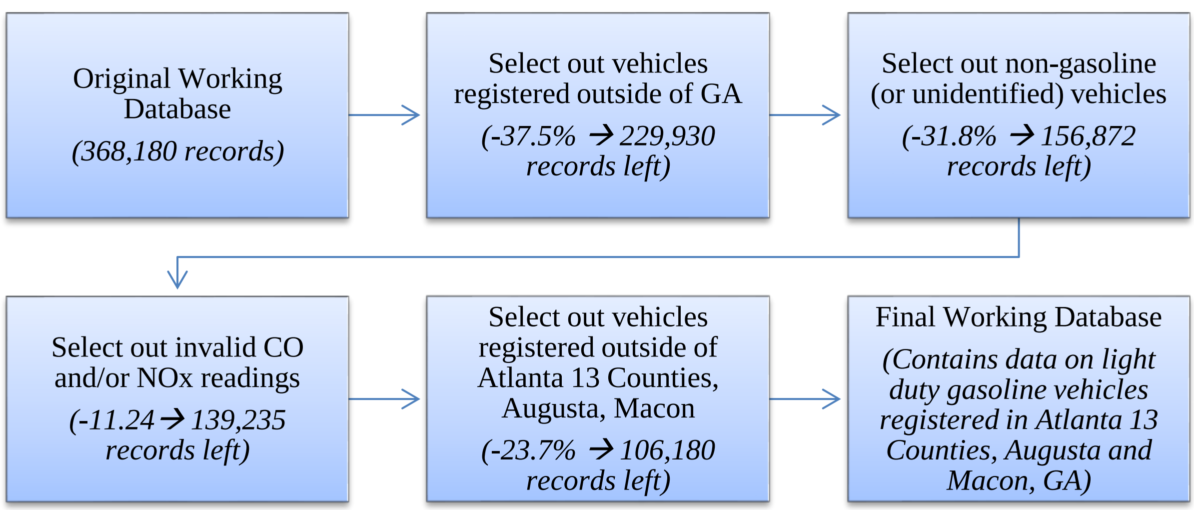

Figure 3.4 Data Reduction.

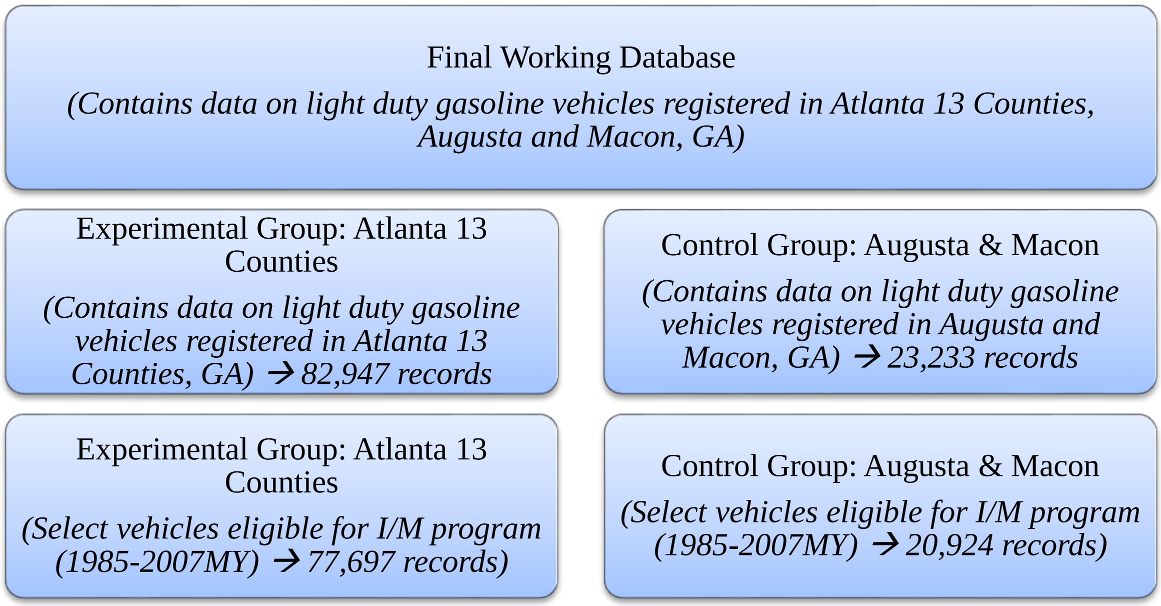

Figure 3.5 Data Split.

The following analyses were conducted in order to fulfill the research objectives stated previously:

1) A series of "null" experiments had been performed on experimental fleet in order to determine fleet characteristics (and their magnitude) that might affect pollution levels across vehicle fleets of the same model year distribution.

If model year is the only parameter affecting exhaust levels then vehicles registered at any designated locales inside Atlanta I/M area should have the same emission distribution. In other words: the light duty fleet registered in Cobb county might be expected to produce the same pollution levels as light duty fleet registered in DeKalb. However, this is not the case (see Figure 3.6). Careful examination of differences in fleet distributions (such as higher percentage of luxury vehicles or vehicles of a certain type) as well as differences in socioeconomic characteristics should result in construction of correction factors that will allow us to compare emissions from any two distinct fleets without noticeable biases.

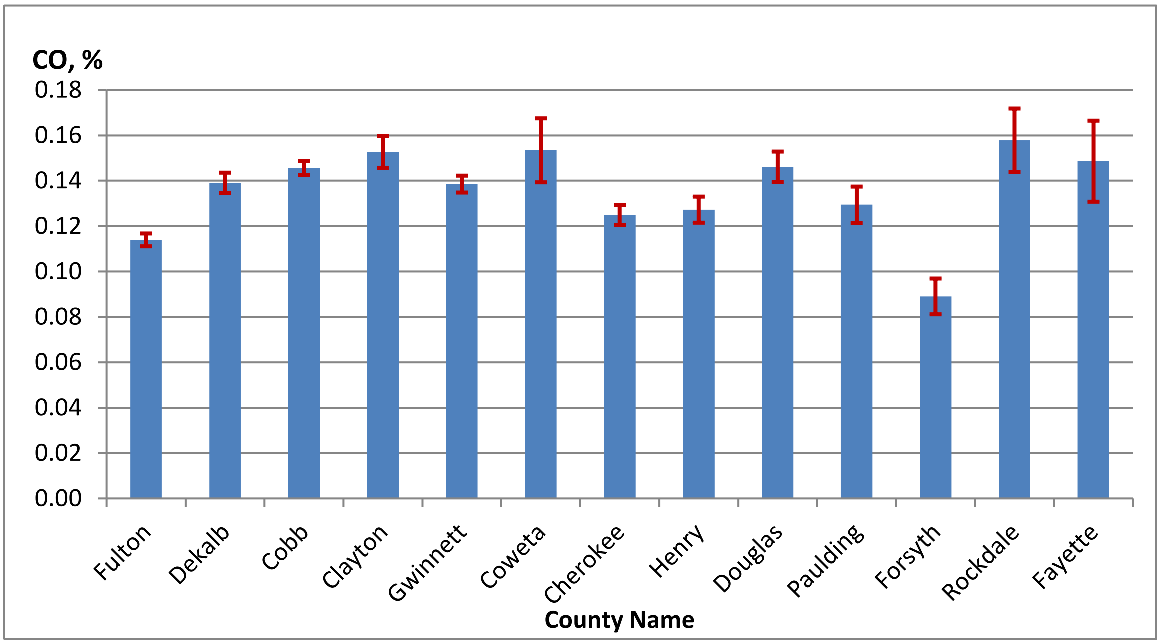

Figure 3.6 Mean CO values across Atlanta thirteen-county I/M area light duty fleets. 2008 measurement year.

2) A "treatment test" was performed on Augusta and Macon reference fleet, and Atlanta 13 counties reference fleet. The presence of the I/M program is considered to be the treatment.

A reference method evaluation with new constructed correction factors for the reference fleet were applied for the latest evaluation of the Atlanta I/M program

3) Evaluation of current reference method.

The results of latest evaluation of the Atlanta I/M program were compared to those produced by the evaluation method developed in this work which includes new correction factors.

CHAPTER 4. ANALYSIS

4.1 Examining Factors Affecting Vehicle Emissions Measurements

4.1.1 Vehicles' Age, Fleet Composition

Fleet composition may play a significant role when it comes to calculating mean emissions for different areas. After all, older vehicles tend to pollute more as do vehicles of certain types. This section of the document presents descriptive statistics on fleet compositions for Atlanta 13 counties I/M area. Data collected in the 2010 calendar year were selected for the purpose of this analysis.

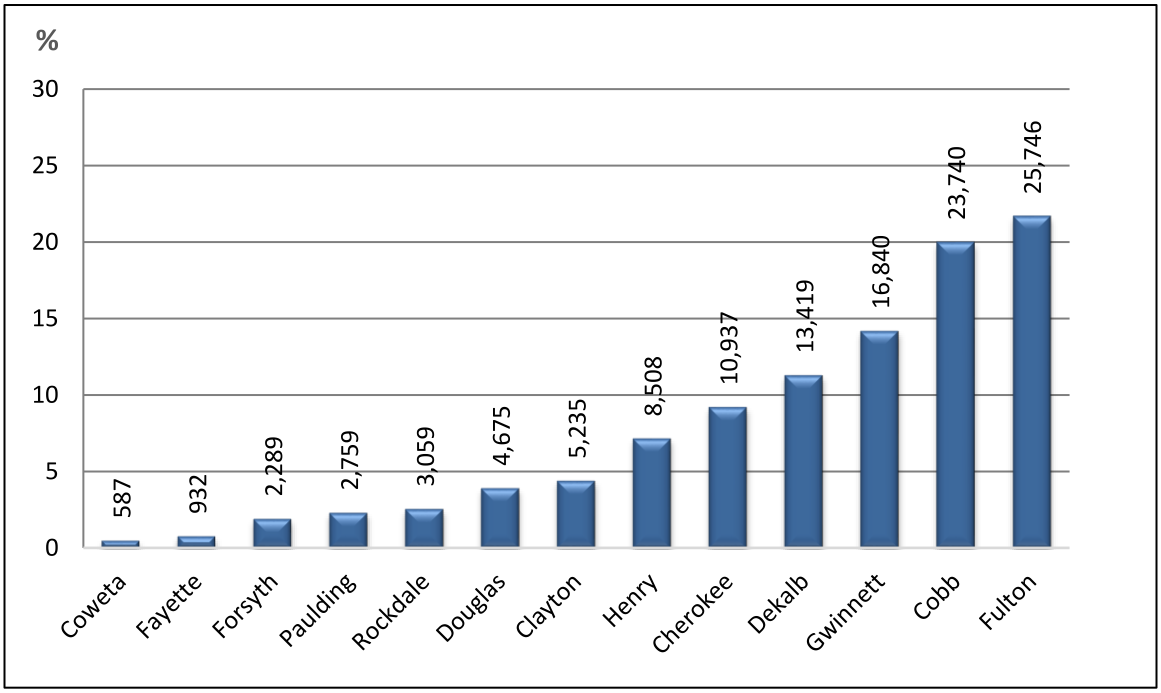

Figure 4.1 Distribution of observed Light Duty (LD) fleet among Atlanta 13 counties.

Figure 4.1 shows that almost 70% of Atlanta I/M fleet is located in four largest counties: DeKalb, Gwinnett, Cobb, and Fulton. Therefore, exhaust emissions from vehicles registered in those counties have a dominant effect on overall emissions.

Figure 4.2 illustrates differences in model year distributions among counties. In 2010 an average car in Atlanta 13 county area was 7.3 years old, with an overall model year mean of 2002.7. The youngest fleet observed was from Forsyth county and the oldest from Clayton county. Error bars represent 95% confidence intervals.

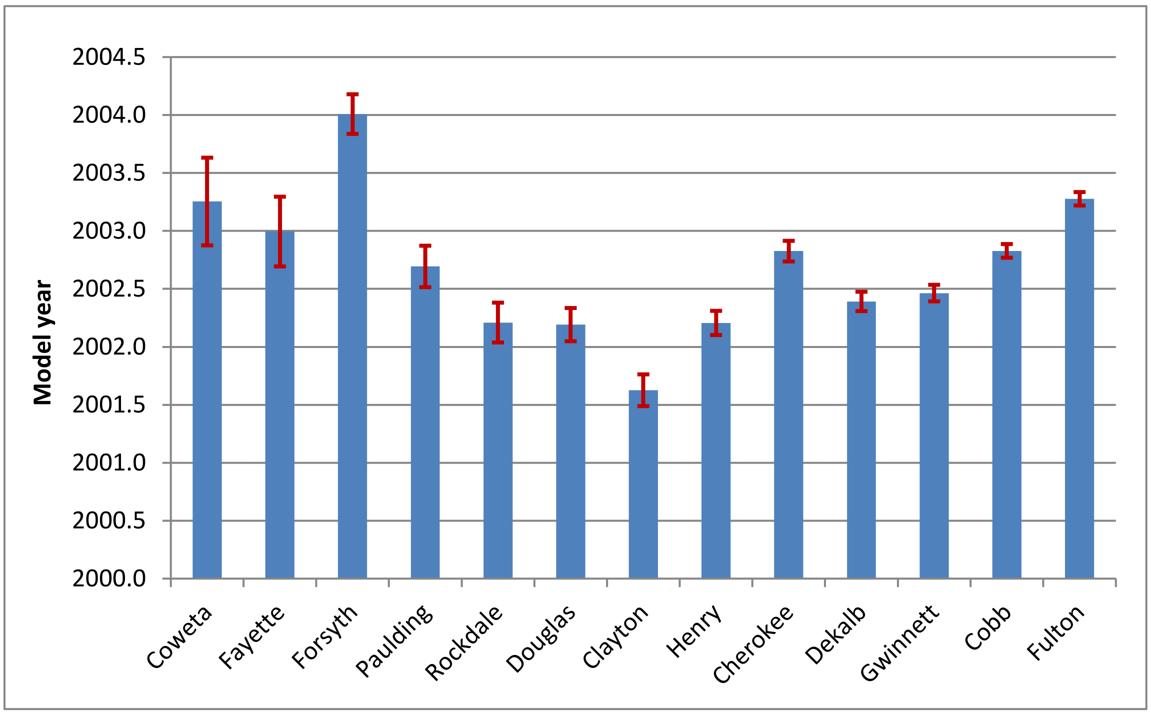

Figure 4.2 Model year distribution among Atlanta 13 I/M counties. 2010 measurement year.

A one-way ANOVA analysis confirms the statement that differences in model year distributions among Atlanta 13 I/M counties are significant (Table 4.1 ANOVA for model year distribution by counties subject to I/M testing.). At the same time certain counties have fleets with similar age distribution: DeKalb's and Gwinnett's fleets have no statistical differences among themselves while Fulton and Cobb have slightly younger fleets.

The effect of model year distribution was thoroughly addressed in a series of previously conducted I/M biennial evaluations (see Appendices A-E). It had been shown that in order to make a fair comparison of emission levels for two or more fleets a normalization by model year is required.

Table 4.1 ANOVA for model year distribution by counties subject to I/M testing.

	Sum of Squares	**DF**	**Mean Square**	**F**	**Sig.**
Between Groups	25,582.90	12	2,131.91	93.47	0.00
Within Groups	2,707,527.45	118,713	22.81		
Total	2,733,110.36	118,725			

Figure 4.3 demonstrates that the fleet composition by vehicle body type among Atlanta 13 Counties is also different. As had been shown in previously conducted I/M evaluations (see Appendices A-E), for thorough assessment of emissions' levels light duty trucks (LDT) and cars should be analyzed separately. However, while this approach could be applied for previous studies, it cannot be used in this one due to low number of measurements collected in some counties.

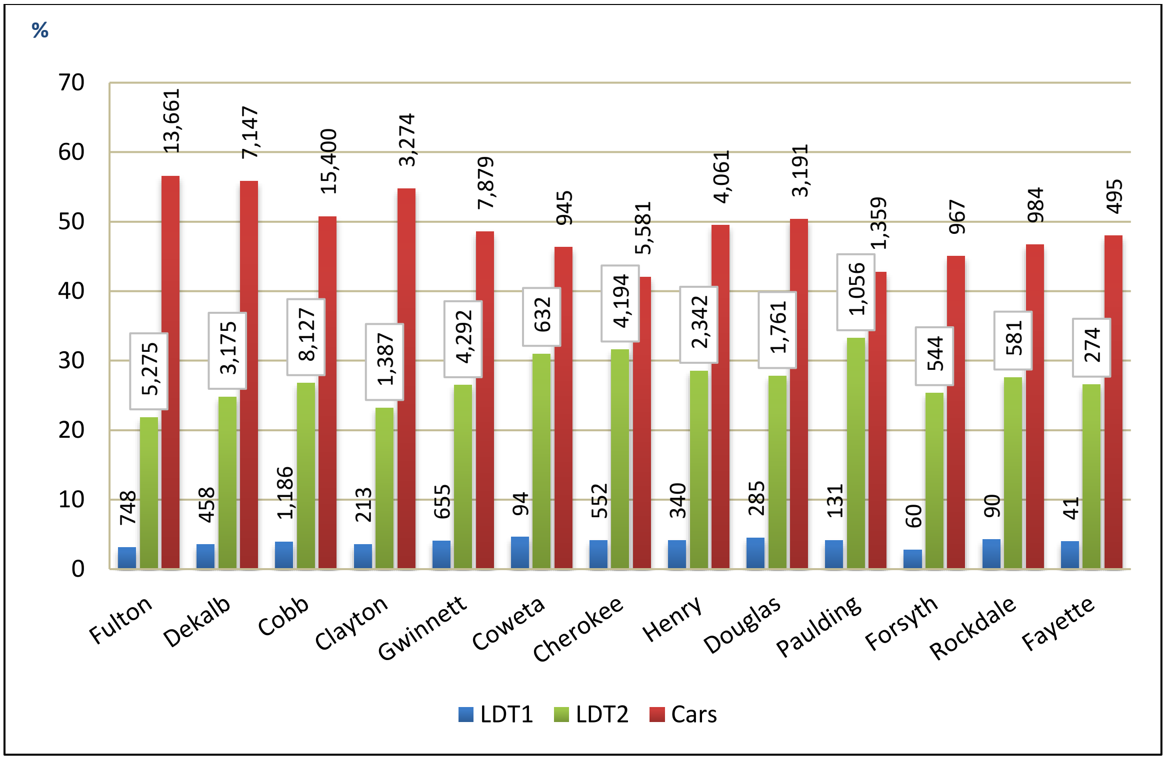

Figure 4.3 Atlanta 13 counties fleet distribution by vehicle body type. 2010 measurement year.

4.1.2 Income, Fleet "Make-Model" Distribution

Figure 4.4 represent income structure among the Atlanta 13 I/M counties. This chart is based on 2011 American Community Survey (ACS) and reflects the December 2009 Office of Management and Budget (OMB) definitions of metropolitan and micropolitan statistical areas thus represent a reasonably accurate picture of socioeconomic structure for the time period observed in this study.

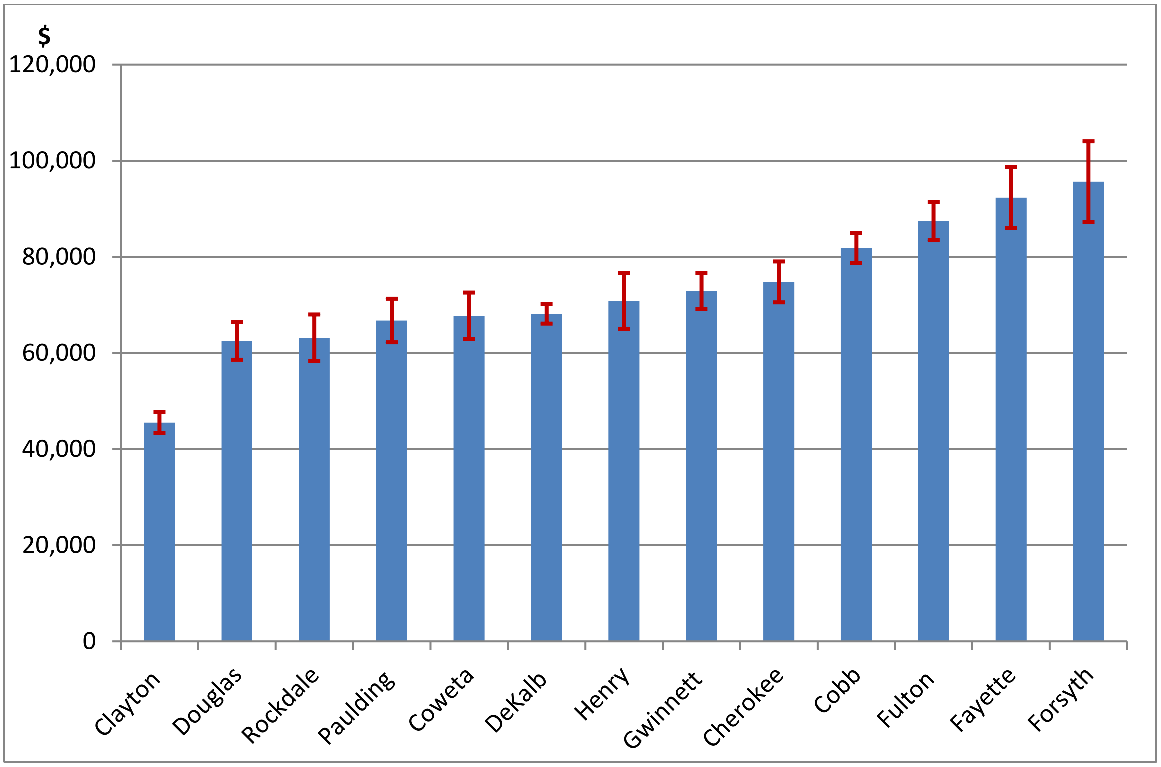

Figure 4.4 Income distribution among Atlanta 13 I/M counties. Source: U.S. Census Bureau, 2011 American Community Survey

Household income of Atlanta 13 county area ranges between 96 and 46 thousand dollars a year. Cobb, Fulton, Fayette and Forsyth form a group of most affluent counties with mean income fluctuating between \$95.6K and \$81.9K. Such a range although seemingly large is not significant from statistical point of view. Clayton has the lowest income level averaging at \$45.5K. Average income of remaining eight counties varies between \$62.5K and \$74.7K with no statistical differences among them.

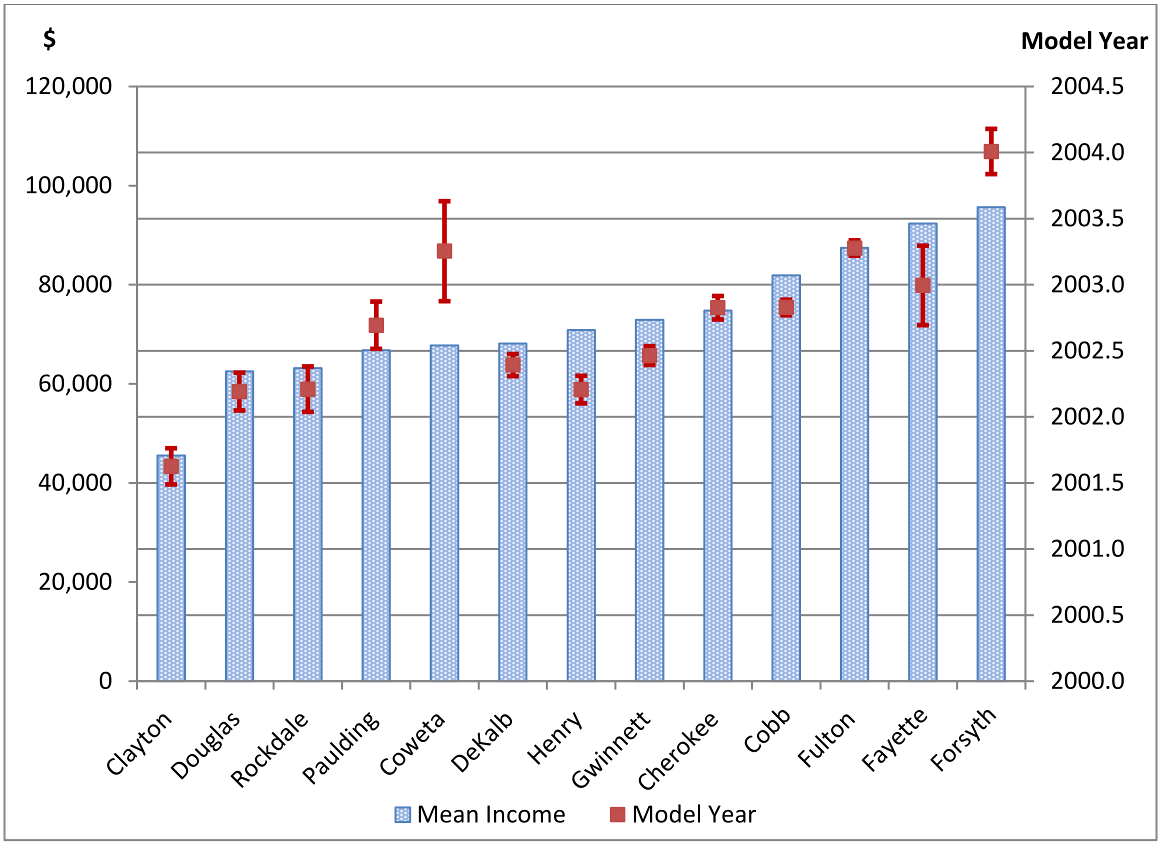

Figure 4.5 Model year distribution vs. income distribution among Atlanta 13 counties.

Figure 4.5 incorporates model year distributions among Atlanta I/M counties into the income distribution. It is natural to assume that average model year will increase with income growth. The figure above generally supports this claim with exception of Paulding and Coweta counties where the fleets seem to be significantly newer than in counties of similar income category. The effect of a counties' model year distribution on emission levels will be addressed later in the text.

Figure **4.6** shows that luxury and imported vehicles are generally more popular in more affluent counties then domestic vehicles. Figure 4.7 links two extreme cases of most wealthy county (Forsyth) and county with lowest income (Clayton). While such deviations in "make-model" distribution may cause differences in emission levels among Atlanta 13 Counties, the data collected does not allow for detailed investigation of this matter.

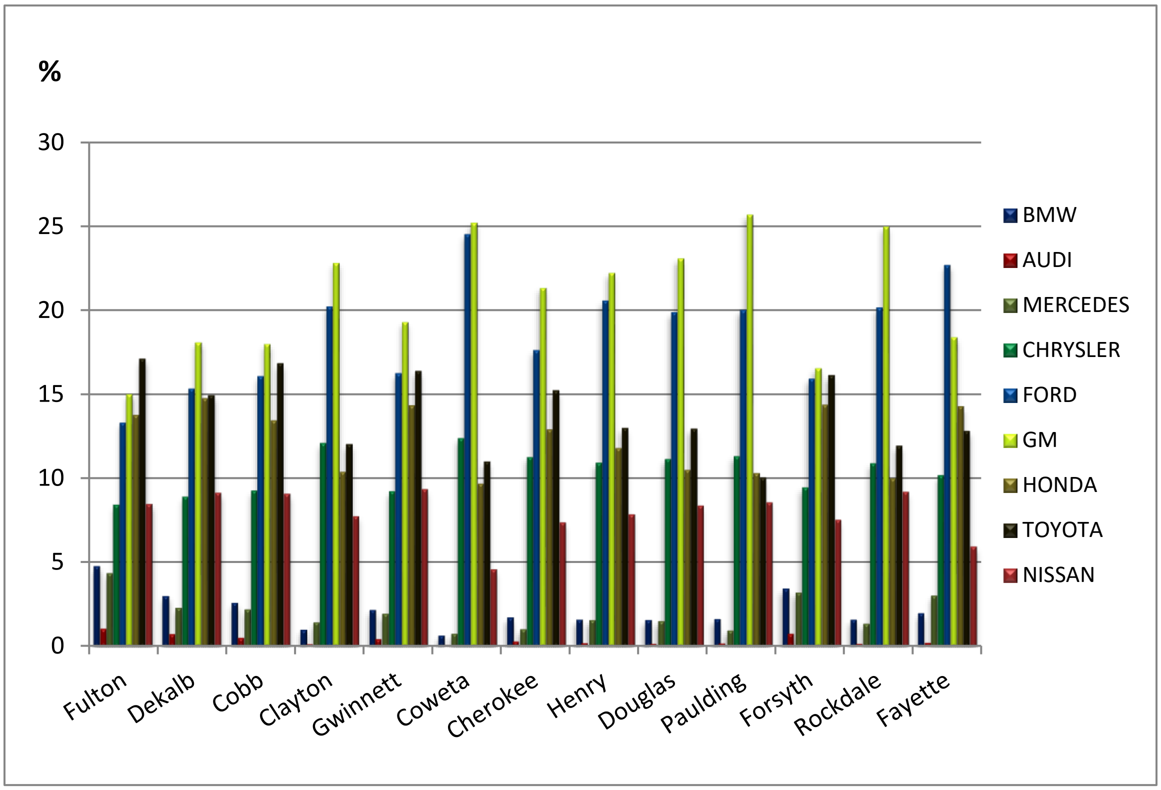

Figure 4.6 Atlanta 13 counties fleet distribution by manufacturer. 2010 measurement year.

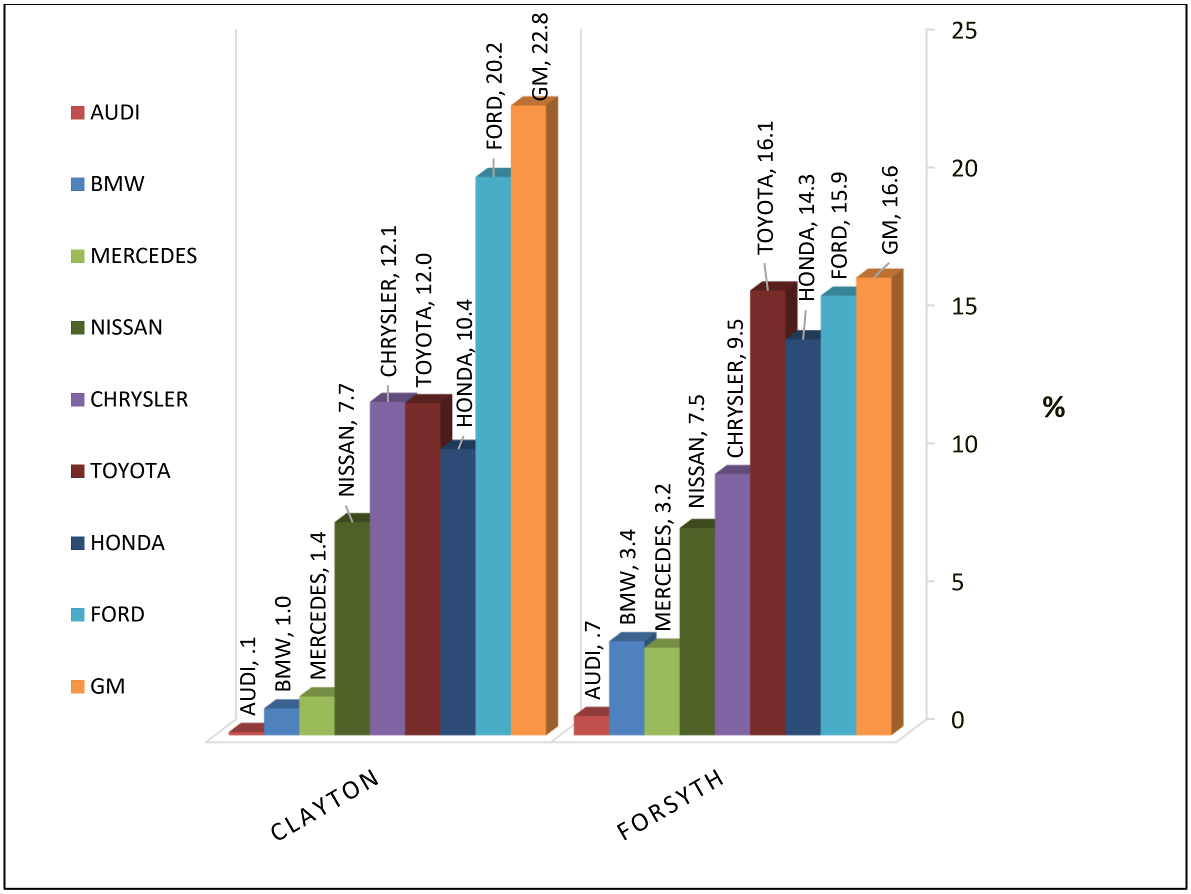

Figure 4.7 Fleet distribution by manufacturer. Selected counties. 2010 measurement year.

4.1.3 Ambient Temperature

Ambient temperatures have been reported to affect automotive emissions since 1966 (Spindt, Dizak, Stewart, & Meyer, 1979). U.S. EPA published a sensitivity analysis for MOVES (Motor Vehicle Emission Simulator) model addressing the effect of temperature and humidity on light duty vehicles' emissions (Choi, Beardsley, Brzezinski, Koupal, & Warila, 2010), stating that "the increase in emissions below 75 degrees is entirely due to the effect of temperature from start emissions" and "for temperatures above 75 degrees, the increase in emissions is due to indirect effect of temperature via air conditioning for CO and NOx". In this study data collection method excludes the effect of start emission. Furthermore, data were collected during the entire year at both experimental and control cites to mitigate temperature and humidity effect. Figure 4.8 illustrates the range of temperatures observed during data collection and numbers of records collected within each temperature block.

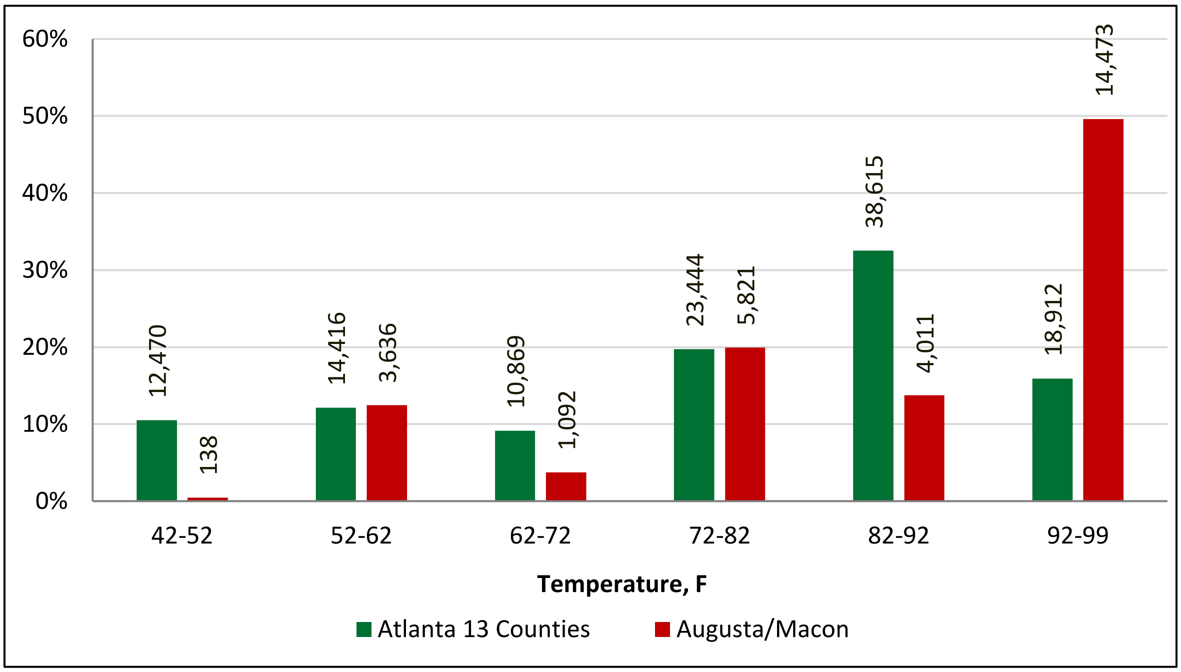

Figure 4.8 Data collection by temperature block. Atlanta 13 counties area vs. Augusta/Macon. 2010 measurement year.

A one-way ANOVA analysis shows that there are no significant differences in CO measurements across temperature blocks for both control (Augusta/Macon) and experimental (Atlanta 13 Counties) areas (Table 4.2 and Table 4.3). The data was controlled for specific VSP range (the reason for VSP control will be discussed in following section of the document).

Table 4.2 ANOVA for CO differences across observed temperature blocks. Atlanta 13 counties. For 6 < VSP < 9 kW/Tonne.

	Sum of Squares	DF	Mean Square	F	Sig.
Between Groups	1.11	5	0.22	1.84	0.10
Within Groups	659.7	5481	0.12		
Total	660.9	5486			

Table 4.3 ANOVA for CO differences across observed temperature blocks. Augusta/Macon. For 6 < VSP < 9 kW/Tonne.

	Sum of Squares	DF	Mean Square	F	Sig.
Between Groups	0.87	5	0.17	0.59	0.70
Within Groups	228.07	781	0.29		
Total	228.94	786			

Figure 4.9 illustrates that there are some differences in NOx readings among temperature blocks for both Augusta/Macon and Atlanta 13 counties areas (the lowest temperature block was excluded from this analysis due to low sample size). For Augusta/Macon those differences seem insignificant due to high variations within the groups. For Atlanta 13 counties area NOx readings for higher temperatures appear significantly lower than for lower temperatures. The use of low sulfur gasoline during summer month could be partially responsible for this effect. Figure 4.10 shows that

differences in NOx readings, however, become insignificant if data is normalized by model year distribution.

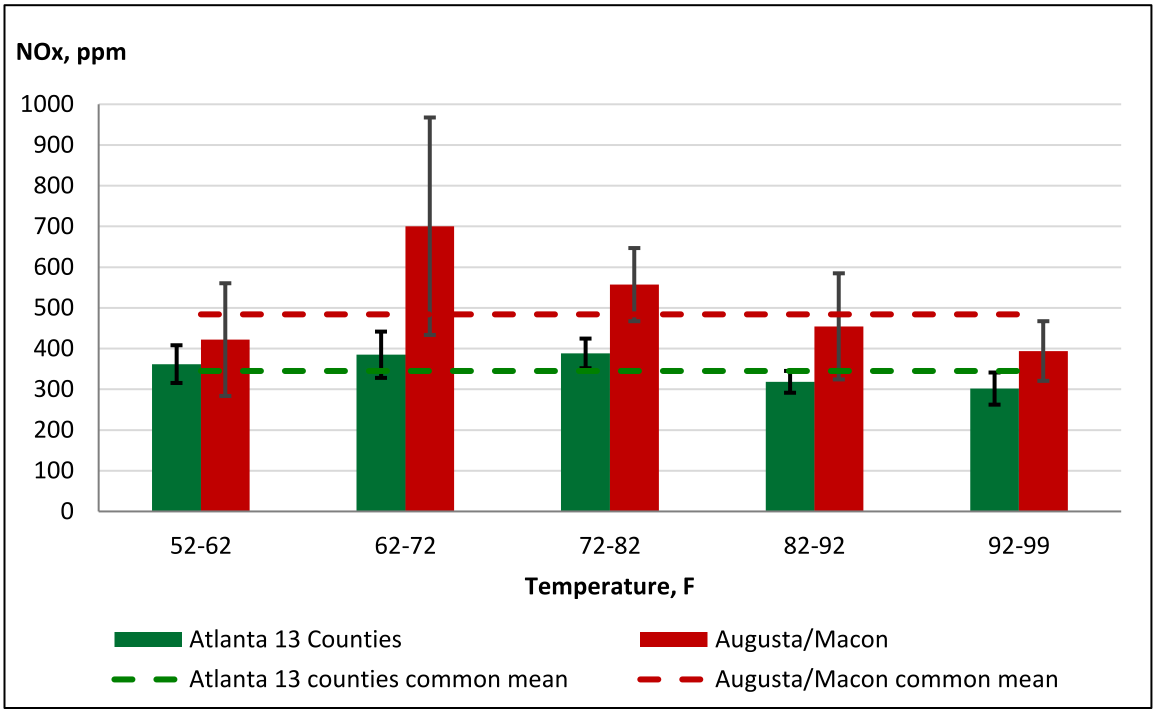

Figure 4.9 Mean NOx levels for Atlanta 13 counties and Augusta/Macon by temperature blocks. For 6 < VSP < 9 kW/Tonne. 2010 Measurement Year.

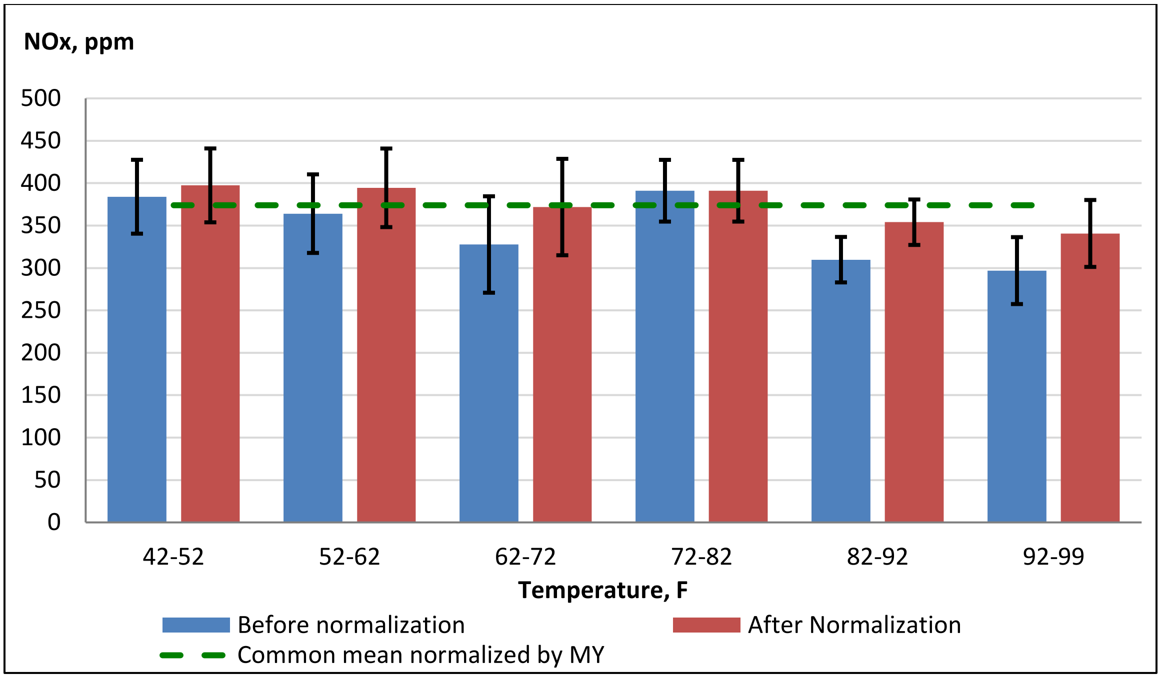

Figure 4.10 Mean NOx levels for Atlanta 13 counties by temperature blocks. For 6 < VSP < 9 kW/Tonne. Normalized to the same model year distribution. 2010 measurement year.

4.1.4 Emission Print and Driving Conditions

It is generally assumed that under same conditions vehicle emission rates would be similar across the Atlanta 13 counties. At least they should not be significantly different. Current evaluation of Atlanta enhanced emission and maintenance program (I/M) is based on this assumption. Vehicles registered in 13 Atlanta I/M counties area undergo annual emission testing and operate on the same regional fuel. However, is it really safe to pool LDV's from entire nonattainment area into one database? Counties that are subject to I/M program have different socioeconomic conditions which leads to differences in the light duty vehicle fleets registered in those counties.

The one-way ANOVA test summarized in Table 4.4 indicates that there are significant differences in average emission rates produced by light duty vehicles registered in the 13 Atlanta I/M counties. *Post Hoc* analysis (Tukey HSD) specifies significant variations in CO and NOx emissions across Atlanta IM area.

Table 4.4 ANOVA for CO and NOx mean levels among Atlanta 13 I/M counties.

		Sum of Squares	**DF**	**Mean Square**	**F**	**Sig.**
CO, %	Between Groups	34.637	12	2.886	10.641	0.00
	Within Groups	32,201.776	118,713	0.271		
	Total	32,236.413	118,725			
NOx, ppm	Between Groups	9.096E11	12	7.580E10	97.159	0.00
	Within Groups	9.262E13	118,713	7.802E8		
	Total	9.353E13	118,725			

Figure 4.11 and Figure 4.12 illustrate 95% confidence intervals on CO and NOx differences between Fulton and other counties that are subject to emission testing. As can be noticed vehicles registered in Fulton counties have significant differences in CO emission levels with Cobb, Clayton, Gwinnett, Henry and Rockdale counties. As for NOx, Fulton's vehicles differentiate from DeKalb, Cobb, Clayton, Cherokee, Henry and Rockdale.

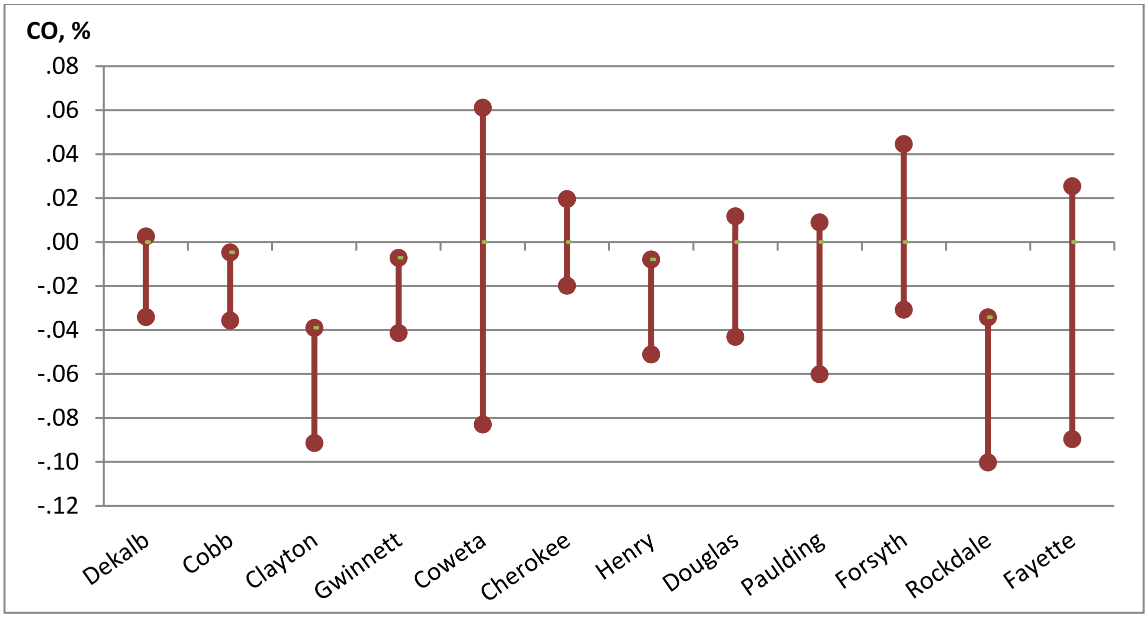

Figure 4.11 95% Confidence intervals for CO emission differences among Fulton and other counties, 2010 measurement year.

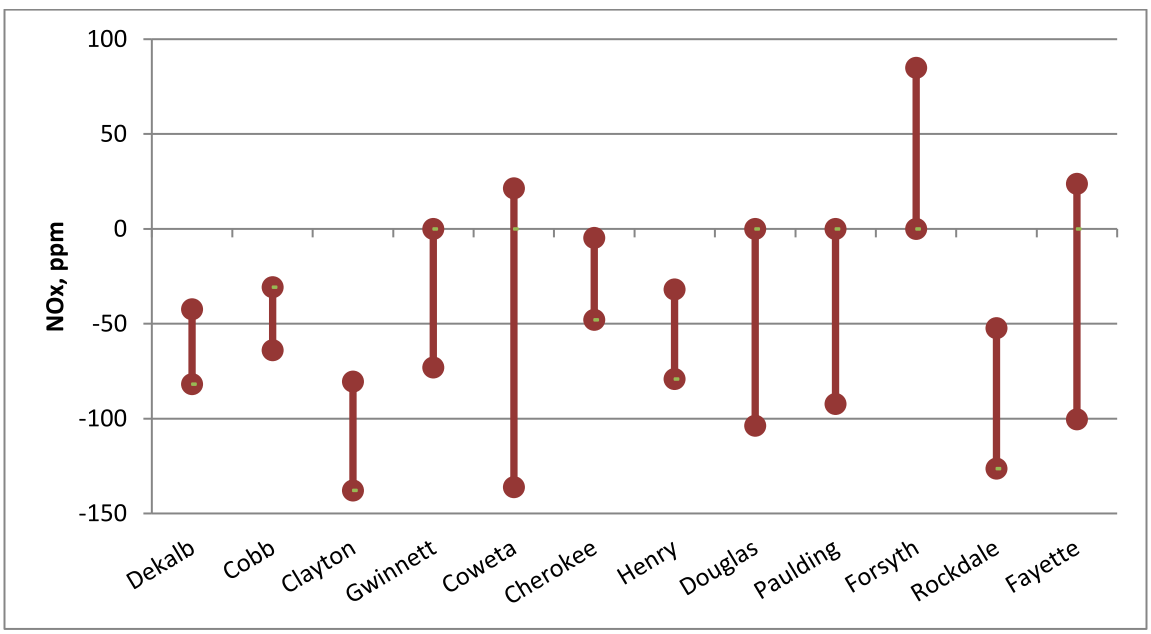

Figure 4.12 95% Confidence intervals for NOx emission differences among Fulton and other counties, 2010 measurement year.

Various reasons might be responsible for the variations in mean emission levels among supposedly uniform 13 counties area; including diversities in vehicles' compositions by model year, type and/or model. Driving conditions also pose its effect on emission levels.

The concept of vehicle-specific power (VSP) has been widely used to describe driving conditions. By dividing the instantaneous vehicle power by the vehicle mass, a parameter named vehicle-specific power was defined by Jiménez-Palacios in 1999, in which VSP was utilized in analyzing the relationship between the field emissions and the vehicle activities. For a typical light-duty vehicle, VSP (in the unit of kW/tonne) could be calculated by using following equation:

$$VSP = v \times [1.1a + 9.81 \times grade(\%) + 0.132] + 0.000302v^3 \tag{2}$$

Where: v = vehicle speed in the unit of m/s; a = vehicle acceleration in the unit of m/s^2; and grade (%) = vehicle vertical rise divided by the slope length, which is generally assumed to be 0.

The later studies (Younglove, 2005) verified that VSP is a convenient single parameter rather than a dual parameter such as speed and acceleration, which has direct physical interpretation of and strong statistical correlations with fuel consumption and emissions. VSP-based approaches have been used in various fuel and emission modeling research, including MOVES (Motor Vehicle Emission Simulator), the U.S. EPA's newest emission modeling tool for estimating emissions from the vehicle fleet in the United States.

It is expected that vehicles driving within same VSP range will not have statistically significant differences in emission rates.

Figure 4.13 and Figure 4.14 support the previous claim showing CO and NOx mean emission levels and 95% confidence intervals for all 13 counties within controlled VSP range.

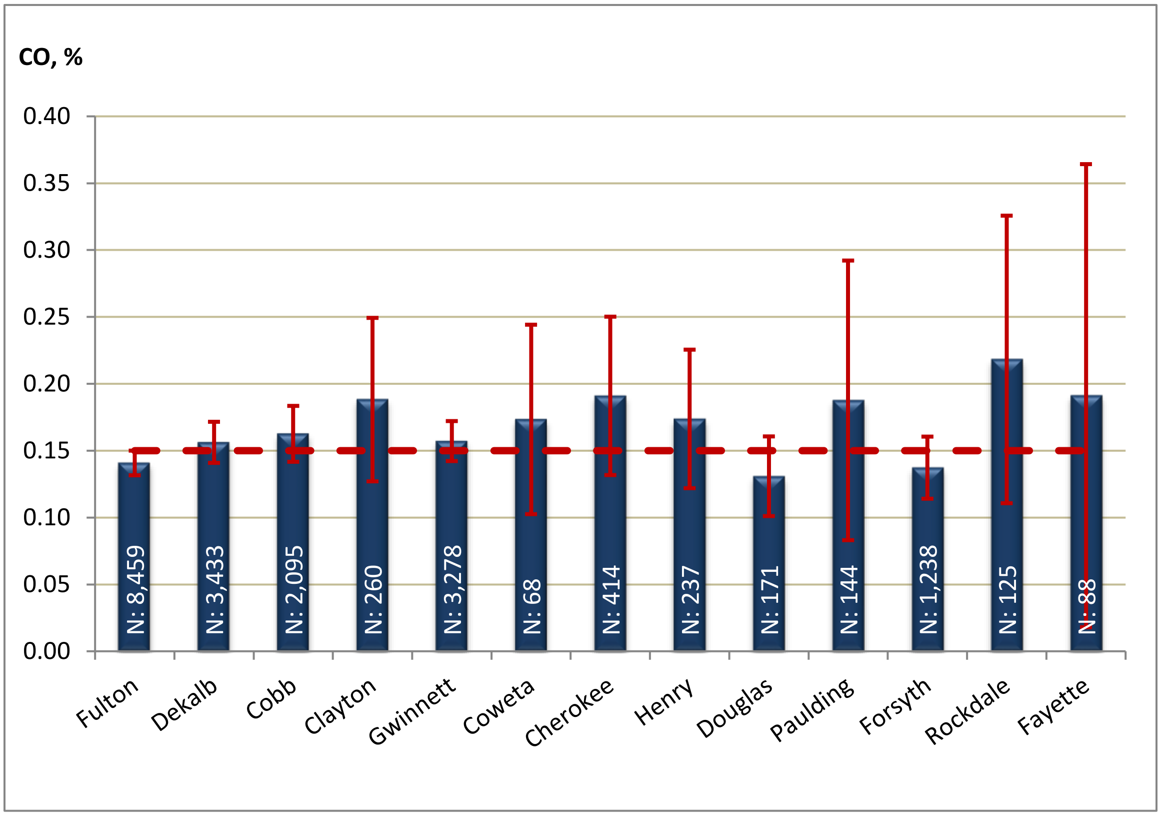

Figure 4.13 Mean CO emissions across Atlanta 13 counties LD fleets for VSP levels less than 10kW/Tonne

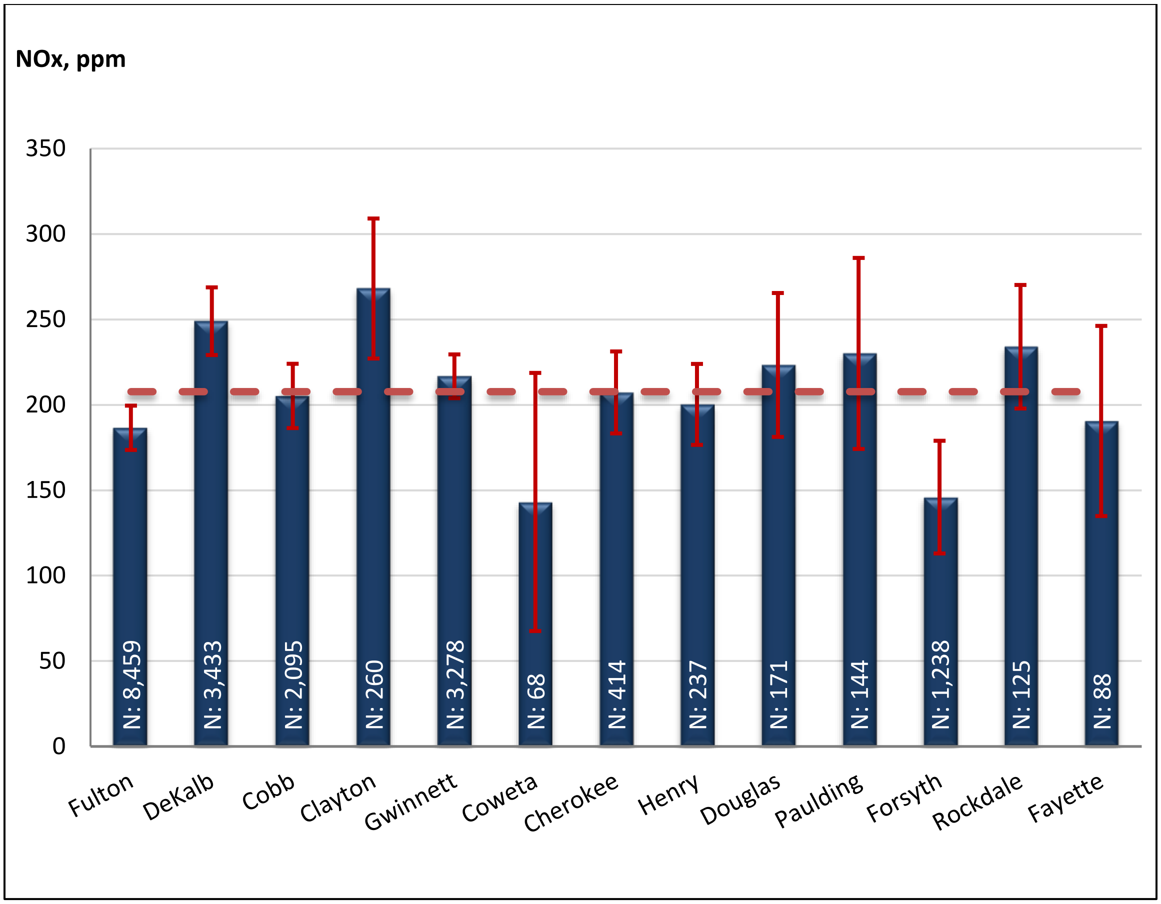

Figure 4.14 Mean NOx emissions across Atlanta 13 counties LD fleets for VSP levels less than 10kW/Tonne

Table 4.5 represents one-way ANOVA results for the same VSP threshold. Low numbers for F statistics demonstrate that there are no statistical differences in CO and NOx emissions among Atlanta 13 counties if vehicles are operated under same driving conditions.

Table 4.5 ANOVA for CO and NOx mean levels among Atlanta 13 I/M counties for VSP less than 10kW/Tonne

		Sum of Squares	DF	Mean Square	F	Sig.
CO, %	Between Groups	3.759E+00	12	0.313	1.532	0.105
	Within Groups	4.090E+03	19,997	0.205		
	Total	4.094E+03	20,009			
NOx, ppm	Between Groups	7.743E+09	12	6.453E+08	1.231	0.254
	Within Groups	1.048E+13	19,997	5.243E+08		
	Total	1.049E+13	20,009			

4.1.5 Comparing Emissions within Same VSP Range

Looking at Figure 4.11 and Figure 4.12 one may ask why there was a difference in average emission levels in a first place. And the answer is: because on the average driving conditions for cars from different counties are not the same. Figure 4.15 illustrates relationship between average emission levels and VSP. Error bars represent a 95% confidence intervals.

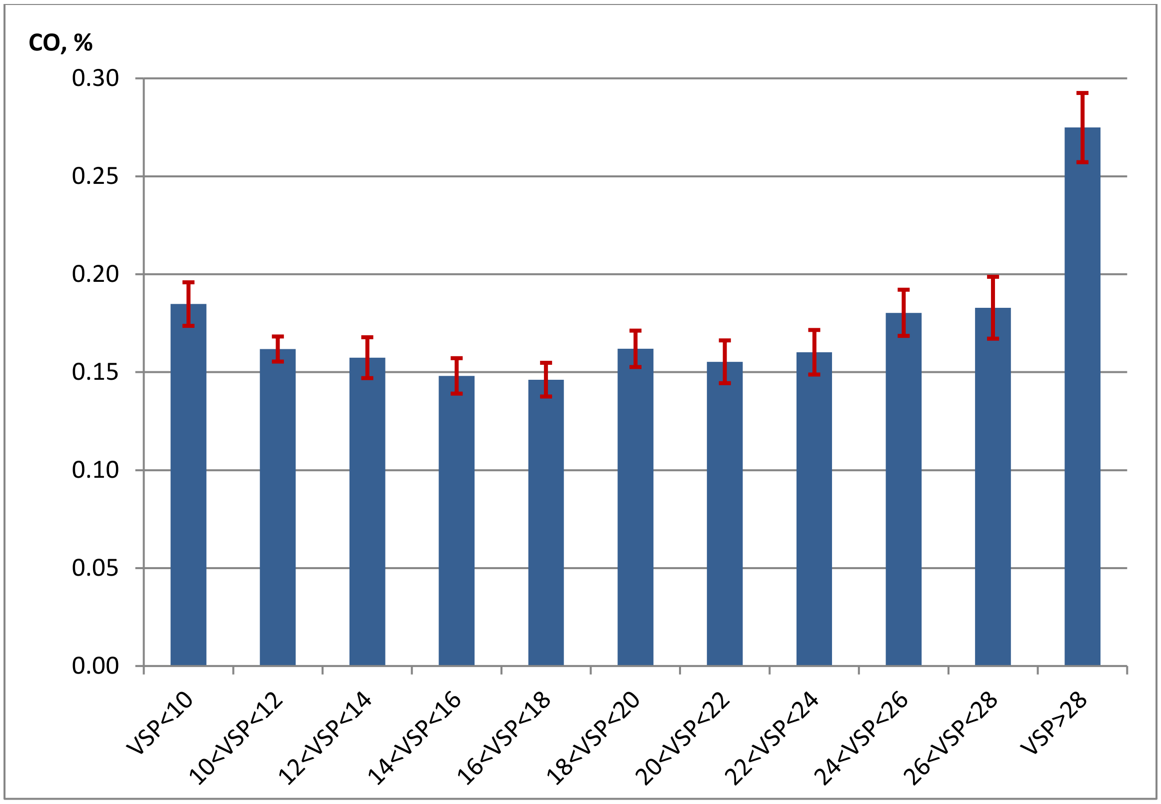

Figure 4.15 Mean CO emissions for different VSP bins

Table 4.6 represents ANOVA results confirming that there are significant differences in mean emission levels for different VSP thresholds. Specifically, vehicles emit more when their VSP levels significantly below or above the average (Figure 4.15). Therefore, to make a valid comparison of vehicle emissions from different counties researchers should: (i) control for VSP levels and (ii) measure emissions across wide range of VSP to ensure that all driving conditions are represented.

Table 4.6 ANOVA for CO and NOx mean levels among range of VSP bins

		Sum of Squares	**DF**	**Mean Square**	**F**	**Sig.**
CO, %	Between Groups	1.01E+02	10	1.01E+01	37.325	.000
	Within Groups	2.69E+04	98,979	2.71E-01		
	Total	2.70E+04	98,989			
NOx, ppm	Between Groups	3.76E+07	10	3.76E+06	12.964	.000
	Within Groups	2.87E+10	98,979	2.90E+05		
	Total	2.88E+10	98,989			

Real-world VSP distributions or detailed speed-acceleration joint distributions that reflect real-world operations are required for estimating vehicle energy consumption and emissions for a particular region. Figure 4.16 shows that in Atlanta and Augusta/Macon areas most vehicles are driven within VSP range between 10 and 28 kW/Tonne.

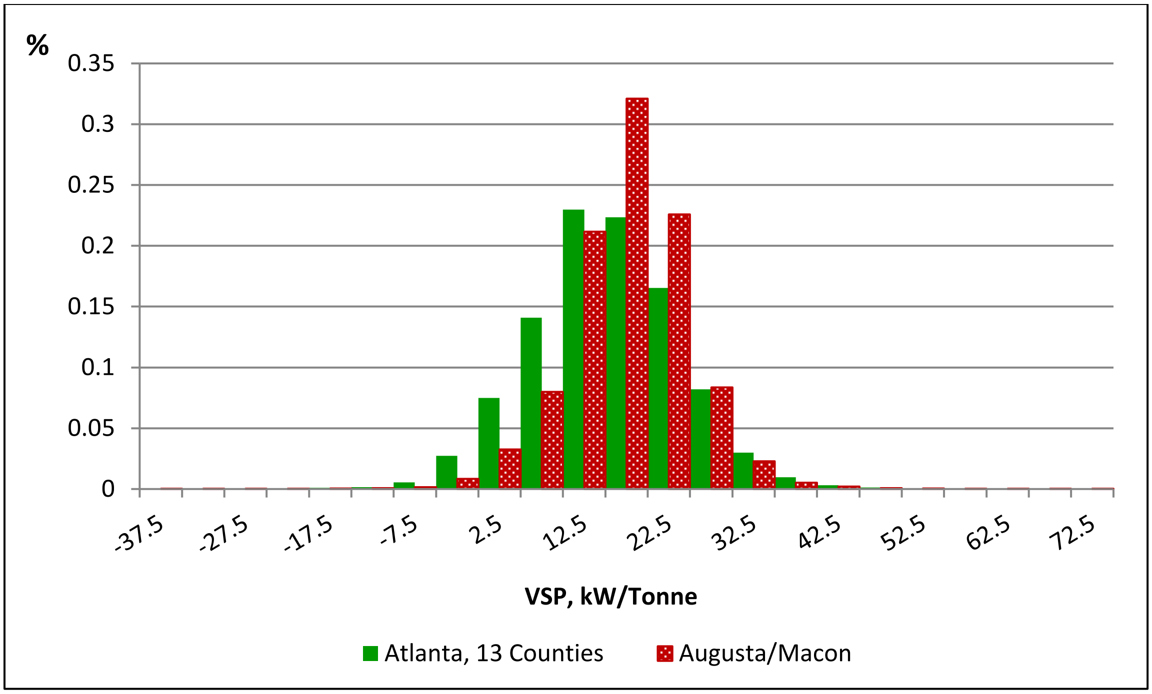

Figure 4.16 VSP distribution in Atlanta 13 county and Augusta/Macon areas.

For the purpose of this analysis, all data were split and then analyzed within eleven VSP bins (see Figure 4.17) that ensures both the similar driving conditions and a reasonable sample size within each bin. It can be seen that VSP distribution is being skewed towards low VSP's numbers especially for the Atlanta 13 county area; however, analysis has shown that more discreet representation of the lower VSP spectrum does not add any additional value to the outcomes of this research.

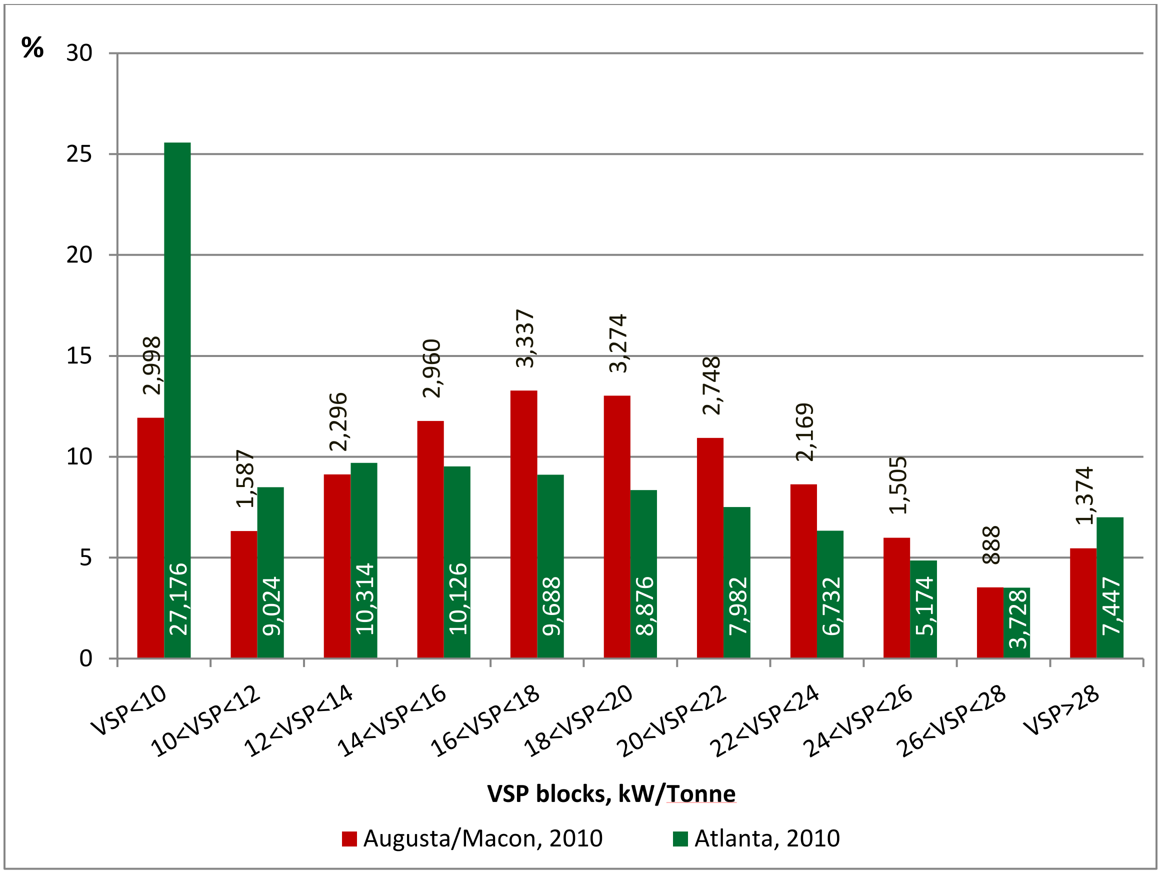

Figure 4.17 VSP bin distributions in Atlanta and Augusta/Macon areas

4.1.6 VSP Distributions within Experimental Sites and Counties

VSP distributions within experimental sites and counties do differ significantly (see Figure 4.18 and Table 4.7). Error bars represent 95% confidence interval levels. It can be explained by the fact that most particular county's fleet is being measured at nearby sites. Therefore, to collect the data that represents all spectrum of driving conditions, it is necessary to measure emissions at several sites that are different by topography and localities.

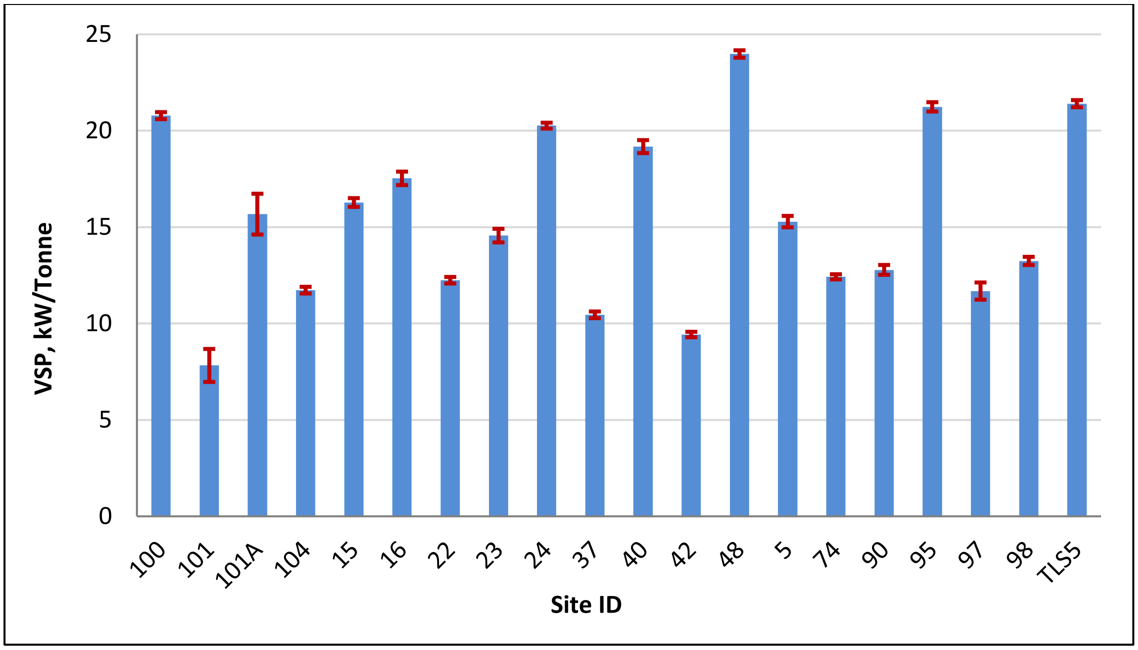

Figure 4.18 Mean VSP levels observed at different measurement site locations

Table 4.7 ANOVA for VSP distribution within Atlanta 13 counties' vehicles

	Sum of Squares	**DF**	**Mean Square**	**F**	**Sig.**
Between Groups	382,924.8	12	31,910	406.23	0.000
Within Groups	7,778,809.1	99,029	78.551		
Total	8,161,733.9	99,041			

4.2 Null Experiment: Comparing Fulton Light Duty Fleet with Gwinnett Light Duty Fleet

In the previous section we extensively discussed factors that can affect vehicles' emission levels and the magnitude of their influence. It has been shown that such features as model year (MY) distribution and VSP distribution may differ significantly even among supposedly uniform Atlanta 13 counties. Variations associated with income levels and temperature are either insignificant or can be considerably mitigated by normalization to a common MY distribution.

The exercise outlined in this section attempts to show that by controlling for MY distribution and VSP blocks all differences in emission readings between selected fleets can be eliminated or significantly reduced. In other words, hypotheses outlined in section "Research Questions" will be tested below.

Two counties, Fulton and Gwinnett were selected for this experiment. Fulton and Gwinnet were chosen based on following criteria:

- A large number of vehicles is registered in both counties providing a significant sample size for the analysis
- I/M program has been present in both counties for the long period of time, therefore, most light duty vehicles had been tested
- Counties are not adjusted to each other
- Income level in Gwinnett county is significantly lower and its fleet is significantly older than in Fulton county (Figure 4.4 and Figure 4.5)

In section 3.4 we hypothesized the following:

H1: There are no significant differences in emission levels among vehicles registered in Atlanta 13 Counties.

H2: Vehicles registered in counties with higher median income levels pollute less than those registered in counties with lower median income.

In order to test these statements for each VSP block emission differences were calculated as follows:

$$\text{Emission Difference} = \left[\sum\nolimits_{\boldsymbol{ij}} \left[\left(O_{GW_{ij}} - O_{F_{ij}}\right)/O_{GW_{ij}}\right]\left(C_{ij}\right)\right]\left(VSP_{ij}\right) \qquad (3)$$

Where: O_{GWij} - emissions observed from Gwinnett vehicles; O_{Fij} - emissions observed from Fulton vehicles; C_{ij} - vehicle fraction from the fleet (based on common MY distribution); VSP_{ij} - fraction of vehicles driven in this VSP block observed by CAFÉ data collection.

Figure 4.19 and Figure 4.20 represent 95% confidence intervals for mean CO and NOx differences for Fulton and Gwinnett light duty (LD) fleets by respective VSP blocks. MOVES model were also run for Atlanta 13 counties area keeping all parameters the same except the presence of I/M program. Looking at the charts it can be concluded that emission differences between Fulton and Gwinnett although not equal to zero are much smaller than ones predicted by MOVES for Atlanta 13 counties assuming no I/M program operates in the area. Given that, results of this analysis support the first hypothesis tested and reject the second one.

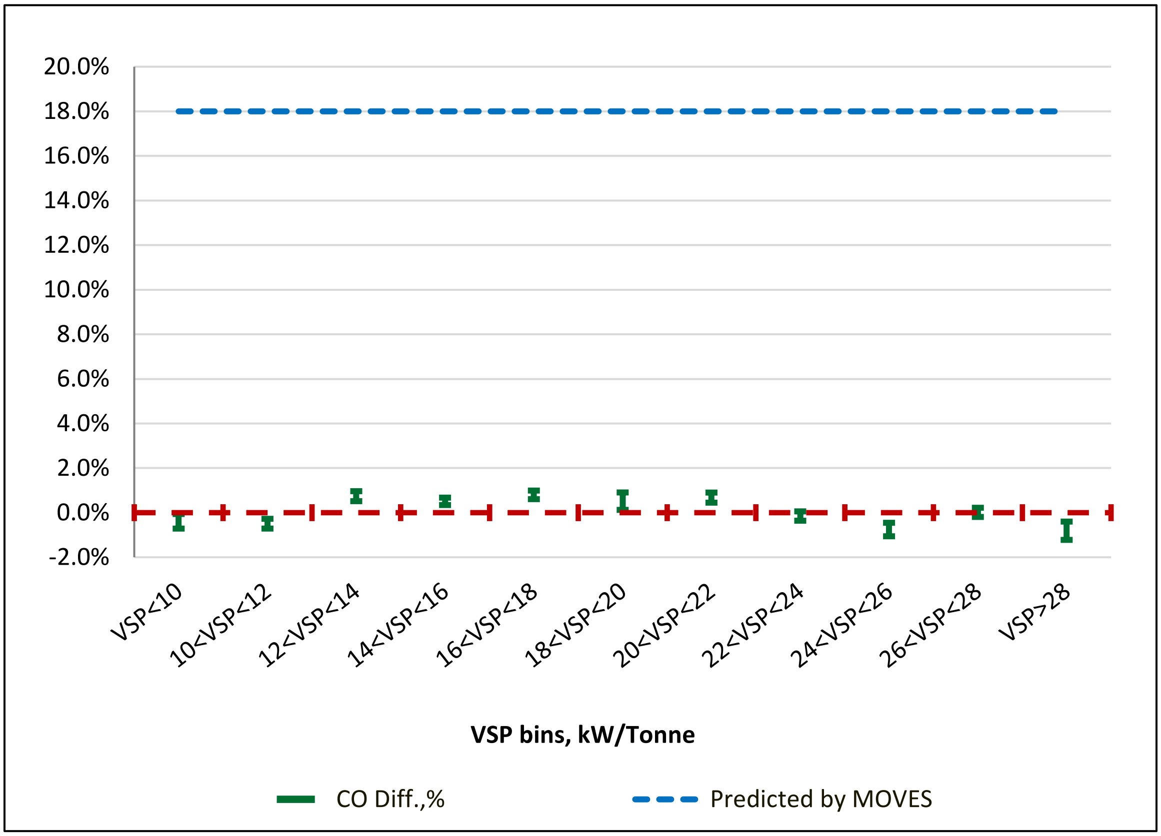

Figure 4.19 95% Confidence intervals for mean CO differences between Fulton and Gwinnet LD fleets comparing with difference predicted by MOVES for Atlanta 13 counties area with and without I/M program. Controlled for MY and VSP. 2010 measurement year. All LD vehicles.

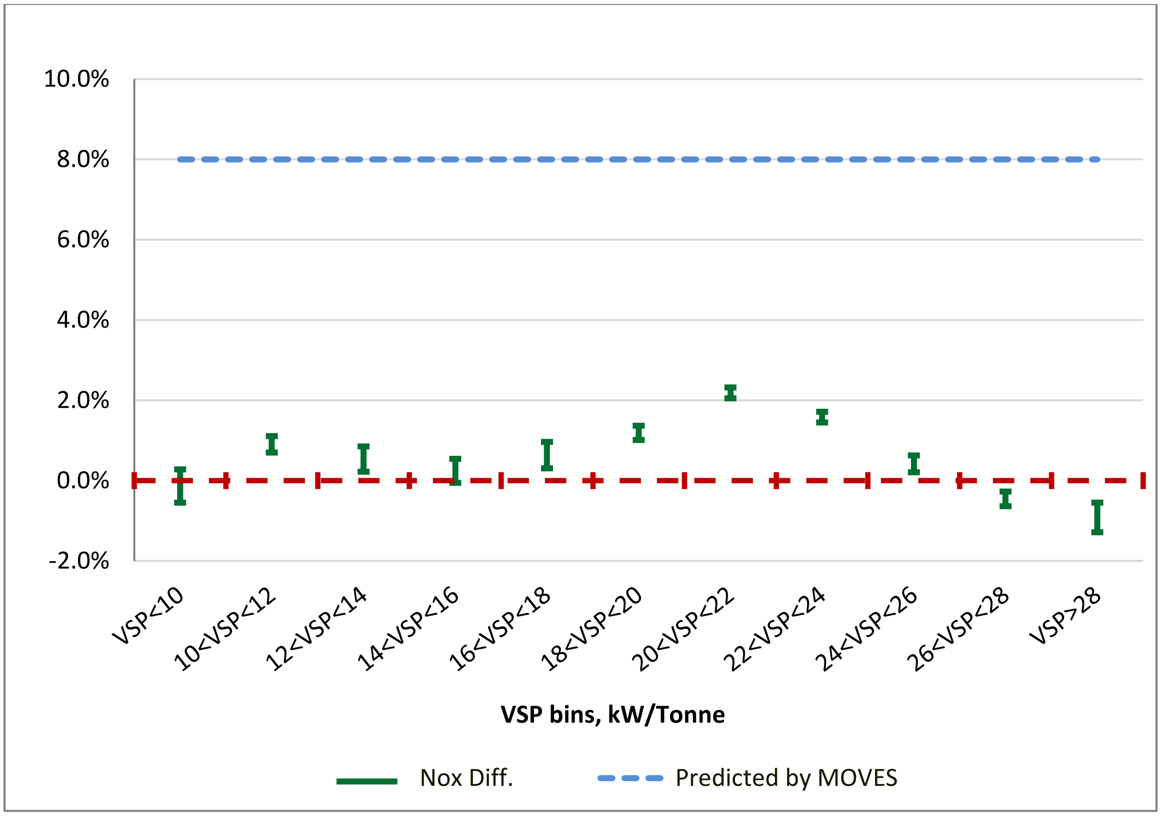

Figure 4.20 95% Confidence intervals for mean NOx differences between Fulton and Gwinnet LD fleets comparing with difference predicted by MOVES for Atlanta 13 counties area with and without I/M program. Controlled for MY and VSP. 2010 measurement year. All LD vehicles.

4.3 Calculating I/M Effectiveness while Controlling for VSP

As was stated in section "Methodology and Analysis Plan", the reference method is essentially a comparison of measurements made at the experimental and reference settings assuming that both settings have the same general characteristics so any differences observed could be attributed to the treatment effect. Applied to Atlanta's I/M evaluation this method compares observed emission differences of inspected and uninspected vehicles assuming that both light duty fleets have the same main characteristics (such as model year distribution, etc.). Results of previous evaluations are thoroughly documented in

Appendices A-E. Socioeconomic parameters as well as driving conditions were not addressed in those assessments, leaving an opportunity for further analysis. This study attempts to close this gap.

Figure 4.13 and Figure 4.14 show that socioeconomic characteristics do not affect emission levels among Atlanta 13 counties when controlled by VSP. Therefore, driving conditions that can be represented by VSP blocks would be the only unaddressed effect on emission levels.

Mean CO and NOx reductions were calculated for each VSP bin using formula:

$$\text{Effectiveness} = \sum_{ij}\left[\left(O_{AM_{ij}} - O_{ATL_{ij}}\right)/O_{AM_{ij}}\right]\left(C_{ij}\right) \tag{4}$$

where: O_{AMij} and O_{ATLij} - average onroad emissions observed for a particular model year for vehicles registered in Augusta/Macon and Atlanta 13 counties respectively; C_{ij} - vehicle fraction from the fleet (based on common MY distribution)

Results of this exercise illustrated by Figure 4.21 and Figure 4.22. Emission reductions for cars 2002 and younger were calculated separately to demonstrate differences in emission reductions between older and newer vehicles and therefore estimate their respective contributions to overall I/M program benefits.

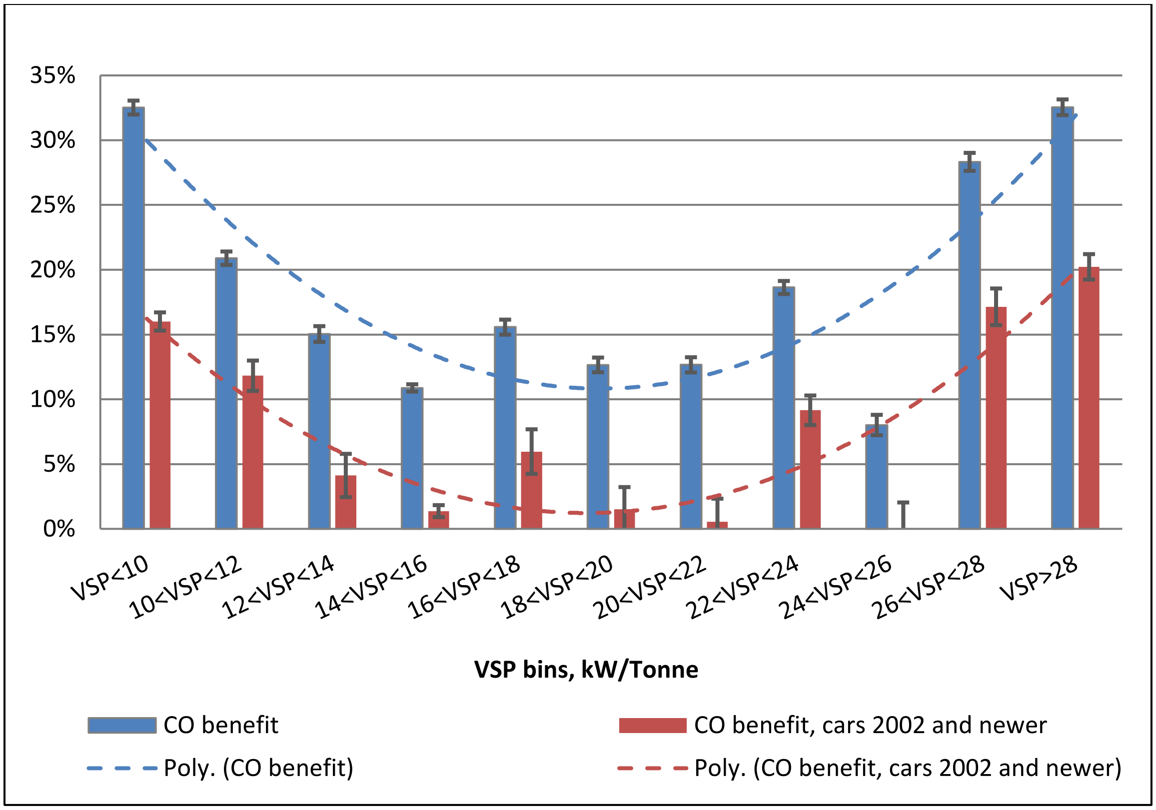

Figure 4.21 Mean CO Emission reductions in percentiles associated with I/M program by VSP bins. 2010 measurement year. Controlled by MY but unweighted by number of cars within the bin.

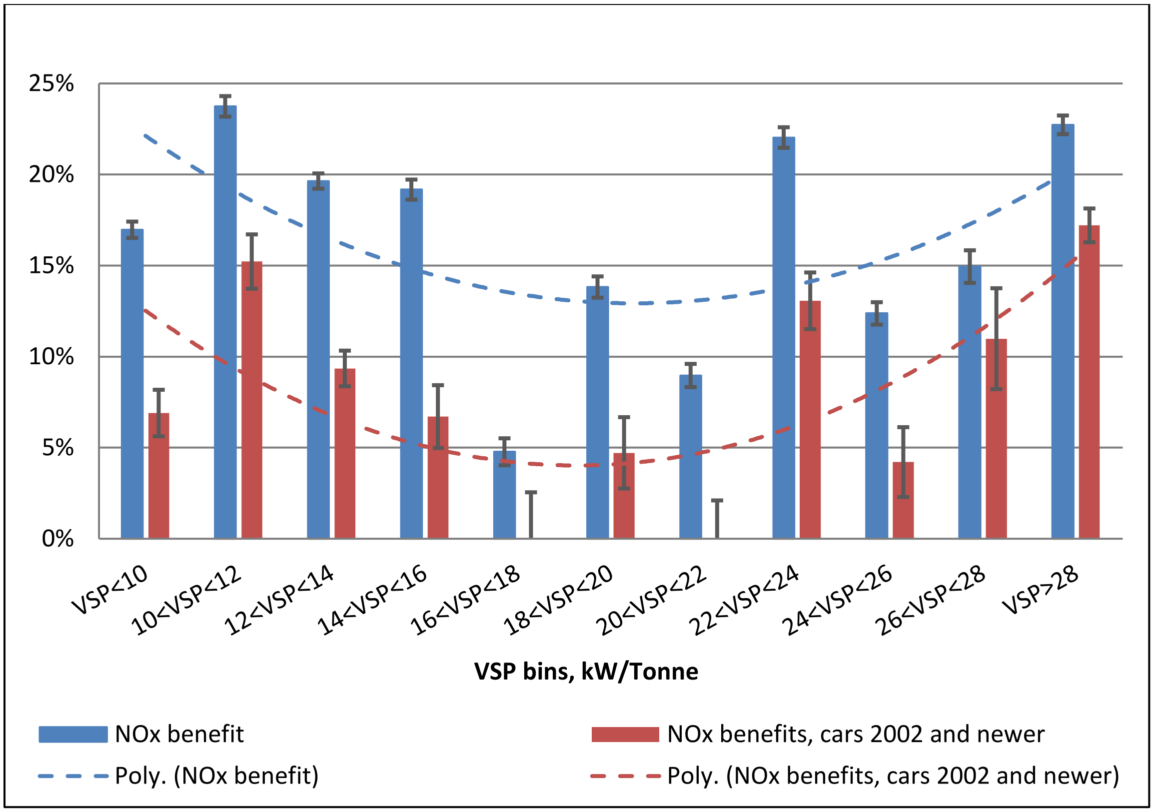

Figure 4.22 Mean NOx Emission reductions in percentiles associated with I/M program by VSP bins. 2010 measurement year. Controlled by MY but unweighted by number of cars within the bin.

Figure 4.23 and Figure 4.24 show that observed CO and NOx emissions in Augusta/Macon area are still higher than those in Atlanta 13 counties area even if controlled by model year and VSP making I/M program the only reason for emission differences.

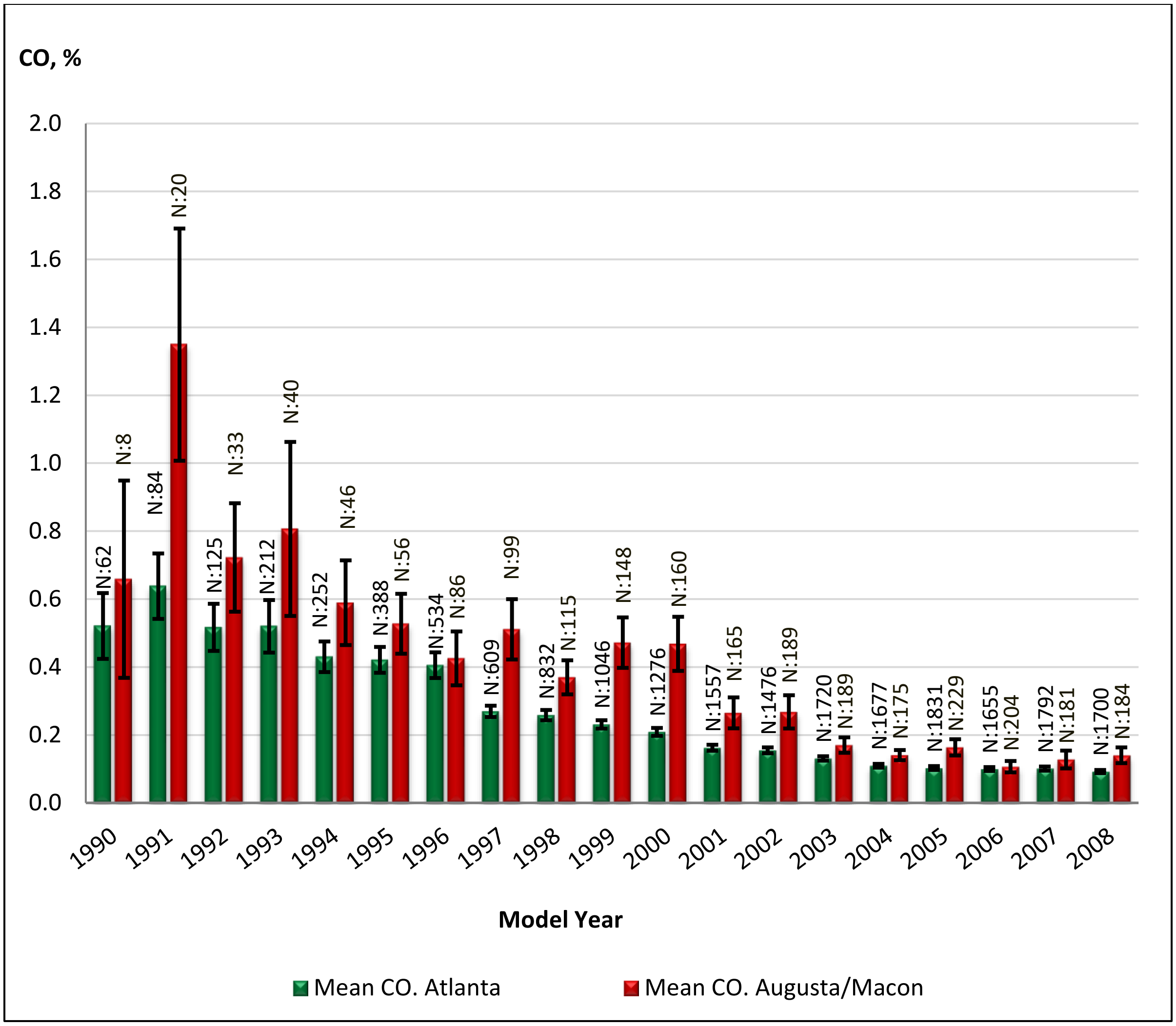

Figure 4.23 CO means by model year. All LD vehicles. Atlanta 13 counties inspected fleet vs. Augusta/Macon fleets. Controlled by MY distribution. Controlled by VSP: VSP < 10kW/Tonne. 2010 Measurement Year.

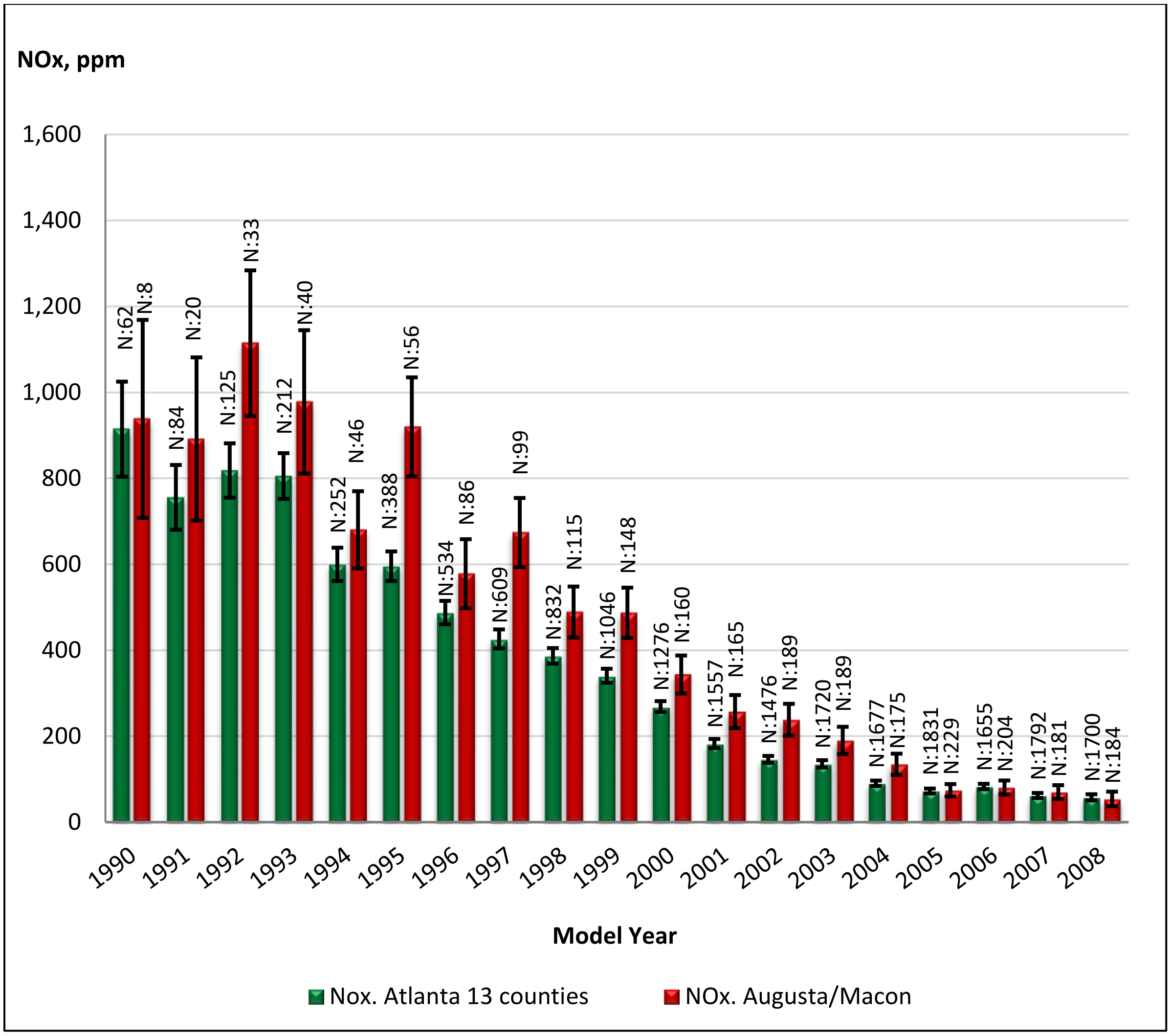

Figure 4.24 NOx means by model year. All LD vehicles. Atlanta 13 counties inspected fleet vs. Augusta/Macon fleets. Controlled by MY distribution. Controlled by VSP: VSP < 10kW/Tonne. 2010 Measurement Year.

Equation 5 represents the formula for estimation of overall emission reductions as follows:

$$\text{Effectiveness} = \sum_{ij} \left[\left(O_{AM_{ij}} - O_{ATL_{ij}} \right) / O_{AM_{ij}} \right] \left(C_{ij} \right) \left(VSP_{ij} \right) \tag{5}$$

where: O_{AMij} and O_{ATLij} - average onroad emissions observed for a particular model year for vehicles registered in Augusta/Macon and Atlanta 13 counties respectively; C_{ij} is the fraction of the Atlanta fleet of that model year and vehicle type observed by CAFE; VSP_{ij} is the fraction of the Atlanta fleet driven in that VSP block observed by CAFÉ

The formula normalizes predicted and observed emissions differences in I/M program and non-I/M program vehicles by model year to the on road fleet fraction and a fraction of vehicles in appropriate VSP bin. This exercise enables the different units of measurement between on road and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form which is important for results' comparison (see Figure 5.2 and Figure 5.3).

The results of this evaluation are summarized in Table 4.8 and Table 4.9. Overall emission reductions for cars 2002 and younger were calculated separately to estimate their respective contributions to overall I/M program benefits. As one can see light duty vehicles manufactured in 2002 and later respectively contribute 42% and 46% to overall CO and NOx emission reductions associated with I/M program although comprise 61% of the inspected fleet.

Table 4.8 Emission reductions associated with I/M program by VSP bins. 2010 measurement year. LDV's all model years.

VSP bin	Reductions Within bins		Weight by VSP bins, %	Weighted Reductions	
	CO	NOx		CO, %	NOx, %
VSP<10	0.3252	0.1697	25.6	8.3165	4.3388
10<VSP<12	0.2089	0.2374	8.5	1.7740	2.0161
12<VSP<14	0.1504	0.1964	9.7	1.4597	1.9061
14<VSP<16	0.1088	0.1917	9.5	1.0363	1.8267
16<VSP<18	0.1557	0.0477	9.1	1.4196	0.4351
18<VSP<20	0.1265	0.1382	8.4	1.0569	1.1543
20<VSP<22	0.1266	0.0896	7.5	0.9509	0.6733
22<VSP<24	0.1863	0.2203	6.3	1.1803	1.3956
24<VSP<26	0.0801	0.1237	4.9	0.3902	0.6024
26<VSP<28	0.2833	0.1494	3.5	0.9938	0.5242
VSP>28	0.3254	0.2273	7.0	2.2805	1.5927
Total Benefits				**20.858**	**16.465**

Table 4.9 Emission reductions associated with I/M program by VSP bins. 2010 measurement year. LDV's 2002 model year and newer.

VSP bin	Reductions Within bins		Weight by VSP bins, %	Weighted Reductions	
	CO	NOx		CO, %	NOx, %
VSP<10	0.1600	0.0690	25.6	4.0932	1.7649
10<VSP<12	0.1181	0.1521	8.5	1.0037	1.2924
12<VSP<14	0.0411	0.0934	9.7	0.3998	0.9074
14<VSP<16	0.0136	0.0670	9.5	0.1303	0.6390
16<VSP<18	0.0596	0	9.1	0.5437	0
18<VSP<20	0.0152	0.0471	8.4	0.1271	0.3939
20<VSP<22	0.0053	0	7.5	0.0403	0
22<VSP<24	0.0915	0.1306	6.3	0.5799	0.8277
24<VSP<26	0	0.0421	4.9	0	0.2048
26<VSP<28	0.1713	0.1098	3.5	0.6012	0.3853
VSP>28	0.2022	0.1720	7.0	1.4173	1.2057
Total Benefits				**8.936**	**7.621**

4.4 Calculating I/M Effectiveness while Controlling by VSP bins defined by MOVES Model

In previous section overall emission reductions where calculated using VSP bins specifically defined for this study. That ensured a granulated enough representation of typical driving conditions for Atlanta 13 counties area and a reasonable sample size within each bin. However, what results one could acquire if VSP bins defined by MOVES model were used? Figure 4.25 and Table 4.10 illustrate results of this exercise. The same approach as defined in section "Calculating I/M Effectiveness while Controlling for VSP" had been applied to calculate overall emission reductions and those within VSP bins. Calculated overall emission reductions associated with I/M program are similar to those obtained in previous section. Slight deviations can be attributed to the error resulting in different vehicles' distribution across the bins.

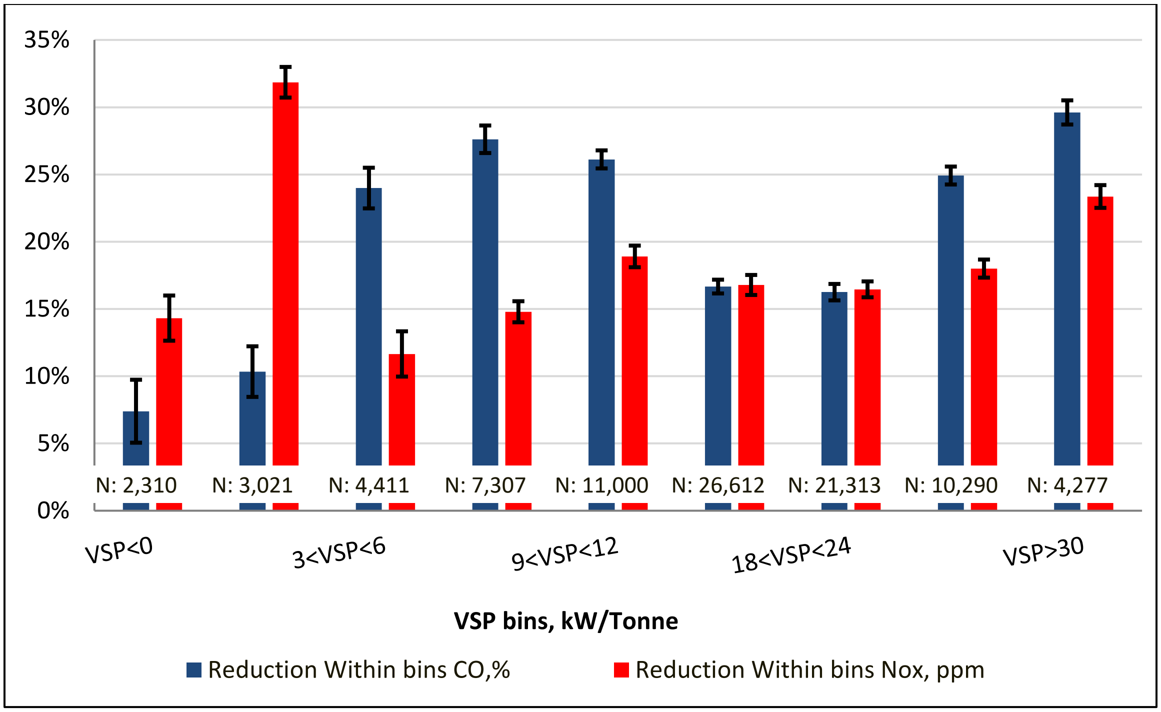

Figure 4.25 Mean CO and NOx reductions in percentiles within MOVES VSP bins. Atlanta 13 counties - Augusta/Macon case. Controlled for MY but unweighted by number of cars within VSP bins. 2010 measurement year. All LD vehicles.

Table 4.10 Emission reductions associated with I/M program by MOVES VSP bins. 2010 measurement year. LDV's all model years.

VSP bin	**Reductions Within bins**		**Weight by VSP bins**	**Weighted Reductions**	
	CO	**NOx**		**CO, %**	**NOx, %**
VSP<0	0.0740	0.1432	3.315	0.2452	0.4747
0<VSP<3	0.1034	0.3185	3.658	0.3783	1.1652
3<VSP<6	0.2399	0.1166	5.053	1.2121	0.5889
6<VSP<9	0.2762	0.1479	7.793	2.1521	1.1527
9<VSP<12	0.2612	0.1891	11.542	3.0148	2.1821
12<VSP<18	0.1667	0.1678	29.480	4.9129	4.9479
18<VSP<24	0.1625	0.1646	24.083	3.9133	3.9637
24<VSP<30	0.2492	0.1800	10.794	2.6900	1.9431
VSP>30	0.2961	0.2336	4.282	1.2680	1.0003
Total Benefits				**19.787**	**17.419**

CHAPTER 5. CONCLUSIONS

Roadside vehicle remote sensing is a common method to evaluate criteria pollutant emissions from vehicle fleets. A popular application of this method is the evaluation of the effectiveness of emissions reduction treatments (e.g. vehicle Inspection and Maintenance (I/M) programs) by comparing the emissions from the treated (experimental) fleet to those of an untreated (reference) fleet. Since no two large vehicle fleets are identical, the reference fleet is produced synthetically by application of correction factors. Current federal guidance assumes that closely located experimental and reference fleets behave similarly if normalized by model year distribution and Vehicle Miles Traveled (VMT).

Atlanta I/M program evaluations outlined in Appendices A-E were performed based on that guidance. As a result, emission differences in the Atlanta inspected fleet and Augusta/Macon uninspected fleets were interpreted as a combined effect of the enhanced I/M program and fuel program assuming that we have controlled for all the differences of those fleets. However, in addition to age and mileage accumulation, emissions from vehicle fleets can be influenced by differences in a variety of other factors including the differences in driving patterns and socioeconomic factors that can impact fleet composition and maintenance trends.

5.1 Overall Results

A series of "null" experiments were conducted to evaluate factors that might affect emissions across vehicle fleets of the same model year distribution. The emission differences among Atlanta 13 counties fleets were addressed and examined to rule out arguments related to different fuel supply, presence of I/M testing and mileage accumulation. However, socioeconomic factors and driving patterns do vary among counties as is shown in Figure 4.4, Figure 4.5 and Table 4.7. Differences in driving patterns should be attributed not so much to motorists' behavior or vehicles' characteristics, but to the measurement site location (see Figure 4.18).

Figure 4.11 and Figure 4.12 illustrate that there are significant differences in average emission rates produced by 13, supposedly uniform, county areas. The present study examined the effect of different driving conditions on emission levels through the concept of vehicle - specific power (VSP) and confirmed that vehicles driven within same VSP range do not have statistically significant differences in emission rates (see Figure 4.13 and Figure 4.14). Another "null" experiment was performed comparing emission rates for Fulton and Gwinnett counties across a range of VSP bins (see Figure 4.19 and Figure 4.20) confirming that there are none or very little differences among fleets driven under the same conditions. Based on this finding a new correction factor was added to evaluation of Atlanta 13 counties I/M program. Socioeconomic factors, as it turns out, have a lesser effect on emission levels, and can be mitigated by normalizing to the same model year distribution.

The overall essential results of work presented can be summarized in following statements:

- This study has shown that quantitative evaluation of emissions control programs is possible with properly selected experimental and reference fleets

- Vehicle's age and VSP have a major effect on emissions
- Differences in emission levels among vehicles registered in 13 Atlanta counties (experimental fleet) can be mitigated by correcting to the same model year distribution and VSP blocks
- If the data is corrected by model year distribution, no evidence was found that vehicles registered in counties with higher median income levels perform better than those registered in counties with lower median income
- Seasonality does not pose a significant effect on emission levels if the data was collected across all seasons and corrected by the model year distribution and VSP blocks
- Contrary to EPA recommendations, I/M evaluations require normalization to VSP, as well as model year and VMT

This study is not without limitations though. The results presented need to be replicated in another geographical area and/or another treatment method (e.g. reformulated gasoline).

The close geographical proximity of the experimental and reference fleets naturally ensure similar environmental conditions and allow for almost concurrent data collection. Therefore, further examination of the seasonality influence on emission readings and overall results should be performed.

Additional in-depth examination of the income influence on observed emissions also has to be conducted. The data at hand did not allow for thorough investigation of different “make-model” distributions on emissions observed. The relationship between income level and vehicles’ maintenance was also not addressed due to the centralized nature of the data available for analysis.

5.2 Main Contributions to the Field

This work represents the first long-term evaluation of an emissions control program (Atlanta Inspection and Maintenance Program) conducted outside of the State of California. It has been demonstrated that a reliable synthetic reference population can be developed and applied in the context of quantitative emission program evaluation. Thorough procedures had been developed to allow other investigators to design and perform similar evaluations. Figure 5.1 summarizes the methodology recommended for such evaluations. The approach outlined in this document can be generalized and applied to a range of emission treatments. In addition, results of this study can be used to validate various emission models, such as MOVES.

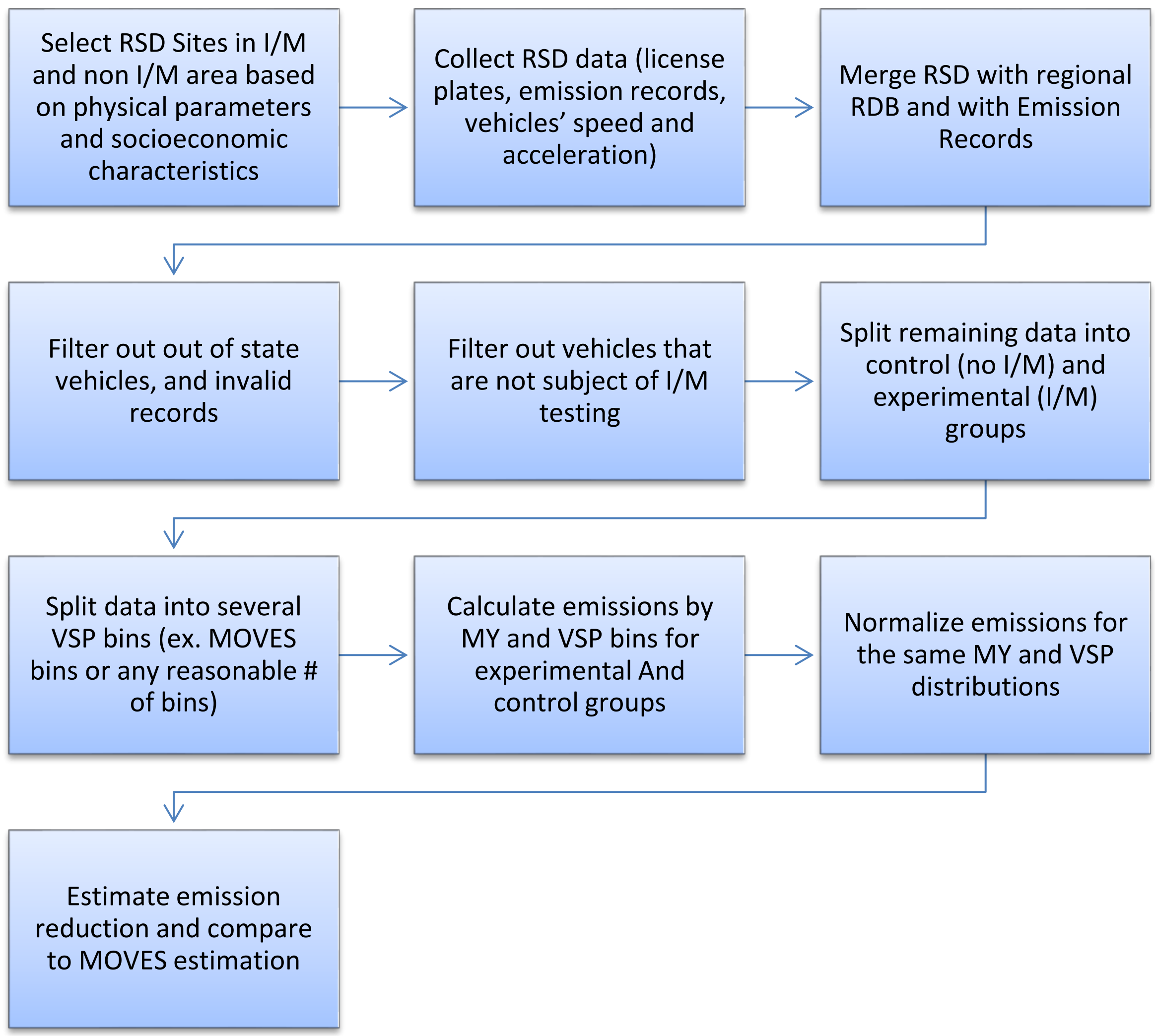

Figure 5.1 Inspection and Maintenance Program Evaluation Methodology

Figure 5.2 and Figure 5.3 represent comparison of CO and NOx reduction attributed to Atlanta 13 counties I/M program calculated using old (EPA recommended) and new methodology. It can be noticed that correction by VSP brings observed emission reduction levels closer to those predicted by MOVES model.

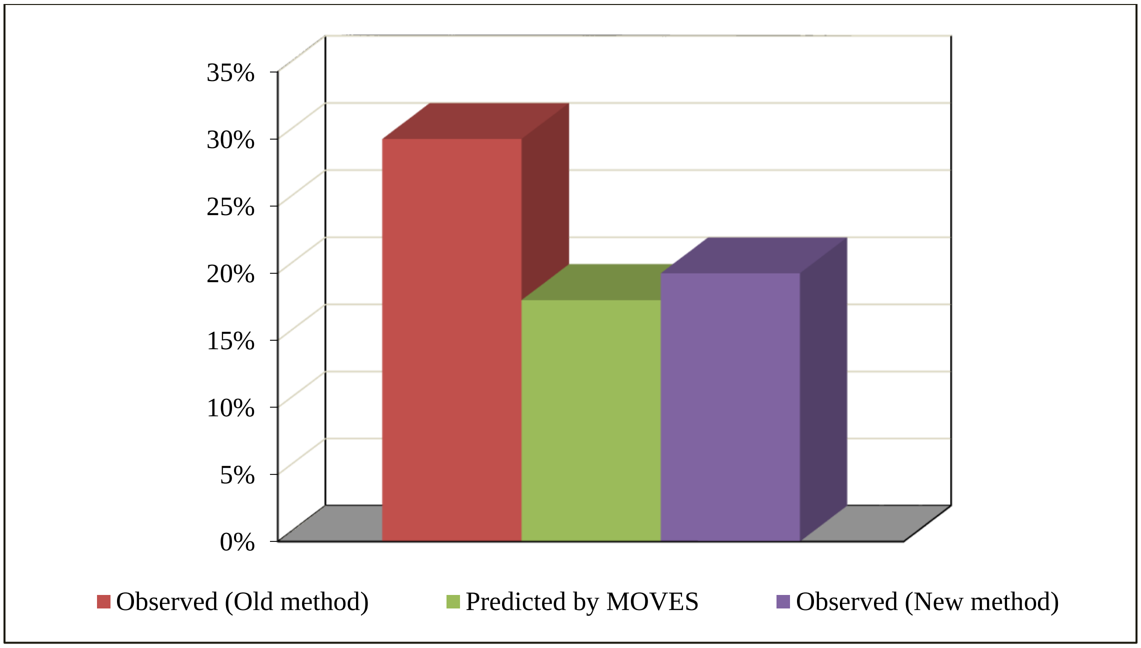

Figure 5.2 CO Reduction in Percentiles. Observed vs. Predicted by MOVES. All Light Duty Vehicles.

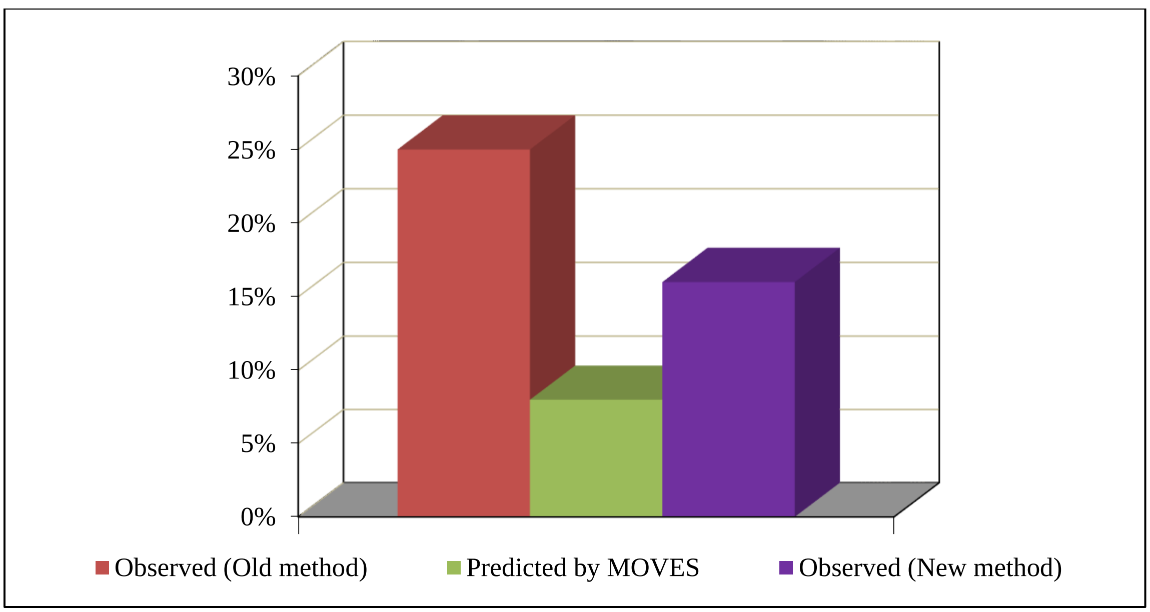

Figure 5.3 NOx Reduction in Percentiles. Observed vs. Predicted by MOVES. All Light Duty Vehicles.

APPENDIX A. BIENNIAL EVALUATION OF THE EMISSIONS REDUCTION EFFECTIVENESS OF THE ATLANTA VEHICLE INSPECTION AND MAINTENANCE PROGRAM FOR 2009-2010

A.1 Introduction

A.1.1 Overview

The 1990 Clean Air Act Amendments (CAAA) made sweeping changes in the scope and stringency of vehicle inspection/maintenance (I/M) programs. Driven by persistent growth in vehicle travel and chronic air pollution in the nation's largest metropolitan areas, the legislation requires "enhanced" I/M programs employ advanced testing technologies and procedures as a way to better ensure the operability of vehicle emission control system.[17] The law also requires biennial evaluation of enhanced I/M programs and on road measurement of inspected fleet emissions, but does not link together the two requirements (CAA Title I §182c3C; CAA Title I §182c3Bi; Title I§182c3Ci).

In the absence of an explicit legislative linkage, the National Research Council has recommended that I/M programs be evaluated using on road emissions data collected by remote sensing devices (RSD) (National Research Council, 2001). RSD uses infrared and ultraviolet technology to measure the emissions of in-use vehicles.[18] The NRC report cited several advantages of RSD data for I/M evaluation. First, RSD is a cost-effective source of evaluation data compared with the higher per-vehicle costs of advanced dynamometer

---

[17] Enhanced I/M programs were required in areas of the United States in serious, severe, or extreme nonattainment of federal ozone standards. Moderate nonattainment areas were required to implement the less rigorous basic I/M programs. Marginal nonattainment areas had no I/M requirement.

[18] Infrared technology is used to measure carbon monoxide and volatile organic compounds. Ultraviolet technology is used to measure nitrogen oxides

testing on a small sample of vehicles, the original evaluation approach recommended by federal regulators. RSD data can also capture trends that cannot be discerned through internal inspection records alone, such as motorists avoiding the program and pre-inspection maintenance behavior. RSD data can also be used for a variety of purposes in addition to I/M evaluation, including mobile source emission inventories, clean-screen programs that exempt low-emission vehicles from subsequent I/M testing, and high-emitter programs that target polluting vehicles for off-cycle inspection and repair.

In response to this growing interest, the U.S. Environmental Protection Agency (EPA) released draft guidance in July 2001 for the use of remote sensing data for I/M program evaluation (U.S. EPA, 2001). The document outlines equipment specifications and measurement procedures along with study design techniques and quality control measures. The document also discusses three methodologies for analyzing remote sensing data to determine I/M program effectiveness. The comprehensive method compares the onroad emissions of the vehicle fleet before and after scheduled I/M testing. The step method compares inspected with uninspected model year emissions during the first year of a new or upgraded I/M program. The reference method compares the emissions of the vehicle fleet located in an I/M area with that of a distantly located non-I/M area.

This paper employs the reference method to evaluate the enhanced I/M program of Atlanta, GA. This major metropolitan area in the southeastern United States is home to

thirteen counties in “serious” nonattainment of the federal ozone standard.[19] The Atlanta enhanced I/M program was implemented in October 1996 in this thirteen-county area, replacing a basic I/M program that had been operating in four of the thirteen counties since the early 1980s. We estimate the effectiveness of the new I/M program by comparing the RSD emissions of a sample of its inspected vehicles with that of a sample of vehicles registered in the Georgia cities of Augusta and Macon. The latter areas have demographics, climate and fleet characteristics similar to Atlanta, but do not operate an I/M program. The emissions difference in the inspected Atlanta and uninspected Augusta/Macon vehicle fleets are then compared with that predicted by the commonly used EPA MOBILE6.2 computer model. Viewing model-predicted emissions differences in inspected and uninspected vehicles as the Atlanta I/M program goal and observed onroad emission differences as actual program performance, we estimate I/M effectiveness as the ratio of these two numbers.

This section provides background on I/M programs, including an overview of I/M program operations, and a history of I/M programs in Atlanta. The second section reviews current enhanced I/M evaluation approaches (including the RSD methods outlined in recent EPA guidance) and their respective strengths and weaknesses. The third section describes

[19] The federal ozone standard is 0.12 ppm averaged on an hourly basis and 0.08 ppm averaged over an eight-hour basis.. Ozone concentration is one of the six National Ambient Air Quality Standards set by EPA to protect public health. (The remainder include carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, particulate matter-10, particular matter-2.5, and lead.) There are five levels of nonattainment for these pollutants, ranging from marginal to extreme, which are determined by the number of times air monitoring stations in an area detect pollutant levels above federal standards during a specific timeperiod.

reference method data sources and methodology. The fourth section reports the results of the reference method. The fifth section discusses the results, with particular attention paid to their caveats. The sixth section presents our conclusions about the reference method results for Atlanta and how it compares with previous I/M evaluations using remote sensing data and reference method.

A.1.2 Vehicle Inspection/Maintenance Programs:

Vehicle inspection/maintenance (I/M) programs seek, first and foremost, to ensure the effectiveness of vehicle emission control systems. The inspection process, which applies to light-duty vehicles of a certain age, involves scheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls.[20] Inspections can be provided by decentralized test-and-repair networks, which allow service stations and automotive repair shops to perform emissions tests and repair failed vehicles, or by centralized test-only networks, in which a limited number of centrally operated facilities perform testing as the sole service. Depending on program design, the test may be performed annually or biennially, while the vehicle idles or is placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of on road driving conditions.

[20] The model year vehicles subject to testing vary across I/M programs.

Motorists must repair failed vehicles, comprising the maintenance component of the program. Vehicles with repair costs above a set amount may qualify for a waiver -- an exemption from further repair and testing -- provided that attempted repairs show some emissions improvement and are not triggered by tampering. Compliance is typically verified through the presence of a vehicle windshield sticker received after passing the test or through the vehicle registration process that requires an emissions certificate.

A.1.3 Inspection/Maintenance In Atlanta:

Atlanta's first I/M program was established in 1981, covering the three ozone nonattainment area counties of Fulton, Cobb and DeKalb. The fast-growing Gwinnett County was added in 1986. The program was implemented through a decentralized test-and-repair network which allowed repair shops, service stations and automobile dealers to perform emission inspections and emissions-related repairs. Testing was originally required for the latest ten model year vehicles, but was expanded in 1986 to include the latest twelve model years. To receive an emissions compliance certificate, cars were required to pass an idle emissions test and an inspection of the catalyst, air pump and fuel inlet restrictor for evidence of tampering. Owners of failing cars that spent more than $50 for repairs qualified for a waiver and an emissions certificate, so long as repairs were not due to tampering and showed some emissions improvement. Owners of cars that failed the tampering inspection were required to obtain repairs to bring their emissions into compliance regardless of cost.

In response to the 1990 Clean Air Act Amendments (CAAA), the Georgia legislature revisited emissions testing in 1992[21]. This legislation enabled the Georgia Department of Natural Resources (GDNR) to upgrade Georgia's I/M program to an "enhanced" program, bringing it into compliance with the 1990 CAAA and new federal I/M federal regulations. This enhanced version of the program received limited implementation in October 1996[22], with emission inspections required only for those vehicles migrating to the Atlanta I/M program area. The new program commenced in January 1997, with biennial emissions testing required of all vehicles from the 1975 model year up until two years of age. The new program also spanned the 13-county nonattainment area, incorporating nine new counties that were not subject to the previous basic I/M program.

After the first two-years of operation several changes had been made to the program: the program began to require vehicles over six years of age to undergo the more rigorous Acceleration Simulation Mode (ASM) testing in October 1998, while in the first two-years period all vehicles were subjects of two-speed idle test (TSI)[23]. The primary difference between ASM and TSI testing is the approximation of real-world driving conditions, i.e., placing the engine under load. While the emissions inspector depresses the accelerator to

---

[21] 1992 Georgia Air Quality Act, Article 2: Motor Vehicle Emissions Inspection and Maintenance Act (OCGA Section 12-9-40 et seq.).

[22] October 1996 was chosen as the soonest possible start-up date after the previous basic I/M program, which operated during a January-to-April vehicle registration "season." Vehicle registration is now conducted year-round in Georgia, as is enhanced emissions testing.

[23] two-speed idle (TSI) testing procedure that measures emissions under idle and a 2500 RPM engine speed

achieve 25 miles per hour (MPH), ASM testing places the vehicle's drive wheels on a treadmill-like dynamometer that applies a 25 percent load on the vehicle engine. The latter approach is more representative of actual driving conditions than an idle test. Vehicles that failed emissions testing were required to be brought into compliance by repair. Owners of covered vehicles in the 13-county ozone nonattainment area were required to show proof of a passing emissions inspection, a waiver, or proof that they qualify for an exemption in order to register their vehicle.

This report concentrates on the seventh two-year period of inspection and maintenance program operation in Atlanta and covers years 2009 and 2010. By this time certain significant changes have been made to the Atlanta enhanced I/M program. The waiver limit had been increased several times. In 2001, testing frequency changed from biennial to annual; the requirement to inspect vehicles back to 1975 model years was replaced with the requirement to inspect the latest 25 model years; and the exemption from testing of the newest two model years was changed to exemption of the newest three model years.

A.2 Enhanced I/M Program Evaluation Methods

Three types of data currently dominate the evaluation of enhanced I/M programs: I/M records, which document the results of each inspection; roadside pullovers, which administer emissions tests to vehicles of randomly selected willing motorists; and remote sensing data, which measures on road vehicle emissions. This section reviews evaluations employing each data type, along with the strengths and weakness of each.

A.2.1 Emissions Inspection Records:

The most common source of biennial evaluation data comes from emissions test records generated by I/M programs.[24] I/M test records provide a cost-effective source of evaluation data because they are routinely generated and easily accessible. Because I/M records cover the entire inspected vehicle population, statistical conclusion validity is generally not an issue: evaluators can control for a variety of vehicle characteristics that influence emissions. The availability of odometer data in most I/M records is also advantageous, enabling evaluators to control for the influence of mileage on emissions. A final advantage stems from inspection/maintenance protocols, which are designed to correlate with the Federal Test Procedure[25] and to facilitate quality control.

However, I/M records suffer from weaknesses that limit their reliability as the sole indicator of program performance. Chief among these is the inability to parcel out fraudulent testing behavior, particularly when inspectors substitute clean-emitting vehicles for unrepaired high-emitting ones on the retest (Wenzel et al., 2000). I/M records may also underestimate program effectiveness by missing pre-inspection maintenance performed by some motorists to lower I/M test failure risks. While it is difficult to quantify the impact of

---

[24] Personal conversation with James Lindner of the U.S. Environmental Protection Agency's Office of Transportation and Air Quality, September 26, 2001. Also see http://www.epa.gov/otaq/epg/progeval.htm.

[25] The Federal Test Procedure is an elaborate testing protocol established in the early 1970s to certify manufacturer compliance with the 1970 Clean Air Act-mandated new vehicle emission standards.

such maintenance, it is expected to yield artificially low baseline emissions and thus underestimate program effectiveness. Generally speaking, these weaknesses speak to the role of I/M records as an internal, not an independent, source of evaluation data.

Evaluations employing I/M records also make tradeoffs between internal validity and representativeness of the data. The inspection process employs highly-controlled conditions to ensure that vehicles are measured under consistent circumstances (e.g., engine stress, vehicle speed, and temperature). While these controls reduce confounding influences on emissions, they represent only a fraction of driving conditions that typify onroad driving. Consequently, the ability to extrapolate I/M test emissions to onroad emissions is limited.

To estimate I/M effectiveness, some evaluations calculate the average emissions difference between the initial and final test scores on failing vehicles and assume that the difference is attributable to the I/M program. Three studies used this approach to evaluate different time periods of the Arizona enhanced I/M program. Two of these studies (Wenzel, 1999 and Glover and Brzezinksi, 1999) estimated a 14 percent reduction in carbon monoxide (CO), a 15 percent reduction in hydrocarbons (HC), and a seven percent reduction in nitrogen oxides (NOx). The third study (Ando, et al, 1999), focusing on repaired vehicles, estimated emission reductions of eight, eight and fourteen percent for CO, HC and NOx.

Sierra Research (1998) also compared initial and final emission results for failed vehicles in AirCare, the Canadian Vancouver/British Columbia emissions testing program.

This study estimated I/M emission reductions of 13 percent CO, 9 percent HC, and 4 percent NOx. Replacing initial emission results of failed inspections with EPA model predictions of an untested fleet's emissions, the researchers estimated 16 percent, 20 percent, and 14 percent emission reductions for CO, HC and NOx. The latter emission differences are thought to be higher than the former because model predictions, as opposed to initial inspection results, are not influenced by pre-inspection maintenance behavior.

The Colorado enhanced I/M program was twice evaluated using inspection records. The first analysis, comparing final test scores for vehicles inspected in 1997 with the new program's first 2,138 initial inspection test scores in 1995, indicated CO emission reductions in the range of 30 to 34 percent (Environ, 1998). The second analysis compared failed vehicles' initial and final inspection results from 1998 that had been converted to Federal Test Procedure scores. The comparison, which normalized repair benefits to the entire inspected fleet, suggested that CO had been reduced by eight percent and HC by six percent, with NOx increasing by one percent (Office of the State Auditor, 1999). While the study results seem contradictory, they cover different timeframes, make divergent assumptions (about deterioration rates, the fate of vehicles with final failures) and employ different measures in estimating I/M effectiveness.

One weakness in attributing before-after emission differences to I/M is the potential for "regression to the mean" emissions behavior, in which a portion of I/M failures will

register lower emissions on the final inspection without repair.[26] This phenomenon is driven by tremendous emissions test-to-test variability, the presence of vehicles with marginally failing emissions, and variance in environmental conditions favorable to test performance. Without verifying repairs, the emissions differences between initial and final test scores may overestimate program effectiveness.

A.2.2 Roadside Emission Inspections:

Used primarily in California, roadside emissions tests are administered with the aid of law enforcement officers who randomly pull vehicles over and ask motorists to voluntarily submit their vehicles to an emissions inspection. Volunteer license plate numbers are then used to query the I/M program database to determine those vehicles with and without an inspection in the past twelve months. Recently inspected and uninspected vehicle emissions are then compared to estimate the emission reductions due to enhanced I/M. Roadside emissions estimates of 1999 enhanced I/M program effectiveness indicate emission reductions of 13 percent for CO, 14 percent for HC, and 6 percent for NO (California Air Resources Board, 2000).

[26] Regression to the mean occurs when two imperfectly correlated measures are compared for a nonrandom sample. The nonrandom sample is typically drawn from high or low scorers on either measure. Regression to the mean occurs when the sample mean moves towards the population mean in the absence of intervention. In the context of I/M evaluation, this means that certain vehicles failing their initial I/M test will score more closely to the mean of the population on the retest, i.e., register passing emissions, without repair. Regression-to-the mean can also occur in vehicles that pass their initial inspection but would fail a subsequent retest.

As with I/M program data, roadside pullovers enable the collection of odometer data for mileage estimates. In contrast with I/M program data, the spontaneity of roadside inspections preclude fraudulent test results that overestimate effectiveness, as well as pre-inspection maintenance behavior that underestimates program effectiveness. However, because roadside emissions tests employ a portable version of official inspection procedures, they sacrifice real-world driving conditions. Furthermore, the approach is costly and generates limited data, requiring as many as four technicians and one law enforcement officer to measure approximately 25 vehicles per day (Wenzel, et al 2000, p. III-8). Self-selection bias is a risk because the test is voluntary and tends to yield a ten percent refusal rate (Wenzel, et al 2000, p. III-8).[27]

A.2.3 Remote Sensing Data from Onroad Vehicles:

A second source of data for evaluating I/M program effectiveness, the one used in this study, is from remote sensing devices (RSD) that measure the emissions of vehicles while they are being driven. The advantage of in-transit measurement is the ability to observe a vehicle's emissions under typical driving conditions, which cannot be as easily captured by traditional controlled emissions testing procedures. Remote sensors can measure a large number of vehicles, an important attribute given the need to control for

[27] The evidence of such bias is mixed. One recent study that used remote sensing to measure the vehicle emissions of refusals and participants alike found no significant difference between the two groups (Wenzel et al, 2000, pg. III-8), while an earlier similar study found that refusal vehicles had 2.5 times the emissions of volunteer vehicles (Stedman, 1994).

tremendous emissions variability due to vehicle type, age, make and model, and emission control technology. A final advantage stems from the unscheduled nature of the measurement, which precludes pre-inspection and fraudulent maintenance behavior that can occur when motorists (as with I/M tests) know when a measurement will occur.

In contrast with the highly controlled parameters of the emissions inspection, the physical circumstances of remote sensing data collection are only approximated through sampling site characteristics (e.g., moderate grades to ensure vehicles operate under only a slight engine load and sampling sites that avoid residential areas to minimize inflated emissions from cold engines). Another drawback is that remote sensors capture a split-second emissions reading that may not reflect a vehicle's typical emissions, making larger samples sizes a requirement to average out random emission fluctuations and to profile emissions aggregated within vehicle type (cars vs. trucks) and model year.

Remote sensing data has been used in three ways to evaluate I/M programs. The first method averages the emissions of vehicles measured before initial and after final I/M testing, with the difference attributed to I/M program effectiveness. Dubbed the "comprehensive method" in recent EPA evaluation guidance, emissions differences can also be generated for various subfleets, such as vehicles initially failing and ultimately passing I/M testing versus failing vehicles that never receiving a final pass. This approach enables a variety of I/M-related analyses, such as deterioration rates of I/M repairs, the influence of pre-I/M repairs on emissions baselines, and a comparison with estimates based on I/M records alone. The major disadvantage to this approach is the enormous volume of onroad data required to measure a representative sample of vehicles before and after I/M

testing. Sample size requirements hinge on the probability of measuring vehicles onroad within a specific timeperiod of I/M testing, a probability that fluctuates with testing frequency and the distribution of sampling throughout the year.

The comprehensive method was used to estimate the effectiveness of the California South Coast Air Basin's enhanced I/M program in 1999 (Wenzel et al., 2000). "Smog Check" I/M records were used to delineate tested from untested vehicles by the existence of an enhanced inspection within the past twelve months.[28] A comparison of these vehicle groups indicates a ten percent reduction in CO, a four percent reduction in HC, and a five percent increase in NOx. An earlier remote sensing study in California in 1996 compared the onroad emissions of 3.5 million vehicles 30 to 90 days before with up to 90 days after their basic I/M test (Klausmeier and Weyn, 1997). For those vehicles that failed their initial smog check and then passed, both CO and HC emission differences registered at 20 percent. Normalizing this result to the entire fleet yielded an estimated nine percent emissions reduction in HC and CO. A third evaluation, of the Arizona enhanced I/M program in 1997, analyzed four million remote sensing measurements on 1.2 million vehicles in the Phoenix I/M area (Wenzel, 1999). The results indicated a seven percent reduction in CO and an 11 percent reduction in HC.

[28] Untested vehicles may have been inspected under the previous basic I/M program more than twelve months ago or they may have had an enhanced inspection after the remote sensing reading.

One weakness in the comprehensive method is the potential seasonal effects that results from the year-round testing required to obtain adequately sized samples. Users of this method have also tended to rely on a few high-volume sites, yielding a large number of repeat vehicles that lower the fraction of unique vehicles that could be reached at a greater number of sites.

A second I/M evaluation approach using remote sensing, known as the Step Method, compares inspected with uninspected vehicles during the first year of a new or upgraded program. The uninspected vehicles comprise an internal control group against which to compare the emission reductions of the inspected vehicles. Because the method applies to the early phases of a new or improved program, it can be used only once to assess program effectiveness.

A remote sensing study of the Colorado Enhanced I/M program compared odd (inspected) and even (uninspected) model year vehicles during the end of the first year of a new biennial enhanced I/M program (Stedman, et al, 1997). At that point, in program history, all odd model year vehicles should have been inspected, whereas all even model year vehicles had no reason to be inspected. This timing rendered even model year vehicles the untested control group against which to compare the odd model year vehicle emissions. The comparison of odd and even model year emissions suggested that Colorado's enhanced I/M program had reduced CO between five and nine percent, while HC and NO showed no improvement.

Three factors limit the generalizability of the Colorado study results to enhanced I/M program effectiveness. Remote sensing took place in a single location, which avoids any confounding socioeconomic or physical influences at different sites but limits generalizability to the overall fleet. Furthermore, vehicles traveling past the remote sensing site were decelerating, which does not represent typical driving conditions and is not the optimal condition for measuring carbon monoxide (Environ, 1998, p. 2-19). A third limitation was that the study measured vehicles transitioning from an annual basic I/M program to an enhanced I/M program, rendering it an evaluation of incremental program effectiveness and not a complete estimate of enhanced I/M program performance.

The research that replicates the original Denver Step Method analysis for a 1997 Atlanta I/M program was made by AQG several years ago. This evaluation was conducted separately for the nine outlying Atlanta counties and the four counties that are closest to the center of the city. The results of the analyses are similar to those found by Stedman et al. in Denver. While the Denver carbon monoxide (CO) weighted program benefit was 6.9% , the Atlanta area CO weighted program benefit is found to be 11.5% and 4.9% for the nine-county and four-county Atlanta areas, respectively. We concluded that the 1997 I/M program change in Atlanta yielded a noteworthy and observable change in fleet emissions (Corley, at al., 2003).

A third approach using remote sensing data (the one used in this study) compares the onroad emissions of vehicles registered in an I/M area to that of vehicles registered in non-I/M areas. The non-I/M area serves as a surrogate untested fleet. The validity of this approach relies on the selection of a non-I/M area comparable in fleet age, a well-documented contributor to vehicle emissions; climate, which can accelerate emission control equipment deterioration; and demographics, which influences the age, quality, and maintenance of the vehicle fleet. This approach was originally applied to the basic I/M program operating in four counties of the thirteen-county Atlanta ozone nonattainment area, with the nine nonattainment counties without I/M comprising the untested fleet. The analysis indicates that car and truck emissions for CO were 15 and ten percent higher, respectively, in the uninspected nine-county fleet than in the inspected four-county basic I/M fleet. The study is limited by its inability to control for differences in mileage and socioeconomic conditions between the two vehicle fleets. The same method was also used to evaluate Atlanta Biennial I/M program by comparing Atlanta 13 counties inspected fleet with Augusta-Macon untested fleet. Assuming that on-road emissions differences represent observed effectiveness and EPA approved model-predicted emissions differences represent effectiveness goals, this study had shown that the Atlanta enhanced I/M program appears to be achieving 83% of its targeted emissions reductions (DeHart-Devis, at al., 2002). Later on that approach became a basis for further biennial evaluations of Atlanta enhanced I/M programs.

A.3 I/M Program Evaluation Components

This study employs an I/M program evaluation method that compares the on road emissions differences *observed* in inspected and uninspected vehicles with the same emissions differences *predicted* by a U.S. Environmental Protection Agency mobile emissions model. The model-predicted emissions difference represents the goal of the I/M program, a reasonable assumption given that states use the model to generate the emission reduction credit received for automobile emissions testing programs. The emissions difference observed in on road inspected and uninspected vehicles is assumed to reflect I/M program performance, an assumption rendered plausible only by the comparability of the inspected and uninspected fleets. We will devote attention in the next section to answering the comparability question.

This section describes the collection of data used in the evaluation. It details the Continuous Atlanta Fleet Evaluation (CAFE), the remote sensing study of on road Georgia vehicles that provides on road emissions data of inspected and uninspected vehicles. The MOBILE6, EPA's recommended emissions model, from which we extracted predicted emission factors, is also discussed. The last section outlines the algorithm that combines data from CAFE and MOBILE6 to generate effectiveness estimates for the Atlanta enhanced I/M program.

A.3.1 On Road Emissions Data:

The Continuous Atlanta Fleet Evaluation (CAFE) provides the on road emissions data used to represent, *inter alia*, Atlanta enhanced I/M program performance. CAFE uses remote sensing devices to measure annually the emissions of approximately 400,000 in-use vehicles in the 13-county I/M program area, as well as two cities located more than 75 miles from Atlanta that do not require vehicle emissions testing.[29] The study is an ongoing

[29] Augusta is located 136 miles east of Atlanta, whereas Macon is 76 miles south of Atlanta.

effort started in 1993 to collect vehicle emissions data for assessing a variety of trends, including fleet turnover, emission control deterioration, and socioeconomic impacts of mobile source control strategies.

RSD measures the emissions of passing vehicles remotely and unobtrusively so motorists are minimally aware of the equipment and do not alter their natural driving behavior. To that end, the remote sensing instrumentation is housed in a van parked on the roadside along with a video camera. An infrared light source and its generator are placed on the opposite side of the road or on the median to create a beam of light that traverses the road. When a passing vehicle breaks the beam, it triggers a measurement of hydrocarbons, carbon monoxide, and nitrogen oxides in the exhaust. Simultaneously, a video camera records the vehicle's license plate, which is automatically scanned into the database of emissions measurements.

After data collection, remote sensing measurements are merged with vehicle registration records using the vehicle license plate. The resulting database allows various characteristics of measured vehicles to be identified, including vehicle identification number,[30] make, model year, and vehicle type. License plates are also linked with inspection/maintenance records to identify vehicles with prior emission inspections.

---

[30] Vehicle identification numbers are 17-digit alphanumeric strings that uniquely identify every vehicle manufactured. When decoded, they provide additional characteristics on vehicles. The VIN-decoded

RSD sampling sites are selected to ensure physically consistent but demographically diverse characteristics. Single straight lines of traffic with an average 35 mile-per-hour velocity are sought to facilitate single vehicle measurements and speeds that maximize measurement opportunities. Driver behavior and driving maneuvers are also observed at each site to ensure that remote sensing measurements would not be biased high by acceleration or low by coasting. Finally, notations are made during the site visits regarding any obvious or suspected diurnal patterns that exist which affect the traffic volume. If distinct variations are found to exist in sites ultimately selected, sampling times are scheduled to account for those diurnal patterns. U.S. Census tract data and traffic count reports inform the selection of different income ranges and land uses.

The remote sensing sites relevant to this study reside within the 13-county Atlanta I/M program area, 12 Atlanta counties without an I/M program but subject to the Atlanta clean fuel program, as well as the Georgia cities of Augusta and Macon that have neither program. The latter locales do not require emissions testing and thus provided an uninspected vehicle fleet to serve as a control group for our previous I/M evaluations. These cities were chosen after a review of census data and registration records revealed them to have characteristics – median household income, population density, and fleet distribution -- most similar to Atlanta than three other Georgia cities considered. But for

---

data of particular relevance to this research are vehicle type (car, truck, multi-purpose vehicle, van) and model year.

the reasons which will be explained later for the present analysis we also used as the reference point the data collected from the vehicles registered in 12 counties that surround Atlanta I/M program area.

According to the state regulation, effective April 1, 1999 the sulfur content of all gasoline supplied in a 25 Atlanta region[31] shall not exceed a seasonal average of 150 ppm (by weight) and, effective April 1, 2001, a per-gallon cap of 500 ppm (by weight)[32]. This rule made vehicle operational conditions in Atlanta 13-county nonattainment area and Augusta-Macon significantly unequal. Since there is no mechanism to separate benefits received from the usage of low sulfur gasoline and emission reductions due to I/M program, the usage of Augusta-Macon fleet as a control group for I/M program evaluation became questionable. In the effort to eliminate the fuel effect, the data collected from the vehicles registered in twelve counties[33] that are not subject to the I/M program but receive the same fuel as Atlanta 13-counties I/M program area have been used in the present analysis as a reference point. The data collected on Augusta-Macon sites represents combined benefits from both I/M and GA fuel programs.

---

[31] 25-county Atlanta region include 13-county I/M program area and 12 additional counties without I/M program: Barrow, Bartow, Butts, Carroll, Dawson, Hall, Haralson, Jackson, Newton, Pickens, Spalding, Walton.

[32] Rules for Air Quality Control Chapter 391-3-1, July 20, 2005. http://www.gaepd.org/Files_PDF/rules/rules_exist/391-3-1.pdf

[33] Eight of these twelve counties are in the Atlanta Metropolitan Statistical Area and thus are considered “Atlanta-area” counties.

A.3.2 Predicted Emission Factors:

We used MOBILE6.2, EPA's recommended computer model for estimation of mobile emission factors, to predict emissions differences in inspected and uninspected vehicles.

A.3.3 Evaluation Algorithm:

We estimated Atlanta enhanced I/M program effectiveness by comparing EPA model-predicted emission differences with observed emission differences in inspected and uninspected vehicles. The comparison yields a percentage that represents the proportion of expected emission reductions actually achieved by the program. The formula for estimating I/M effectiveness is as follows:

$$\text{Effectiveness} = \frac{\sum_{ij}[(O_{n_{ij}} - O_{m_{ij}})/O_{n_{ij}}](P_{n_{ij}})(C_{ij})(VMT_{ij})}{\sum_{ij}(P_{n_{ij}} - P_{m_{ij}})(C_{ij})(VMT_{ij})}$$

where: O_m and O_n are the average onroad emissions observed for a particular model year and vehicle type for I/M program and non-program vehicles, respectively; P_m and P_n are the model-estimated emission factors for I/M program and non-program vehicles for a given model year a vehicle type; C_{ij} is the fraction of the Atlanta fleet of that model year and vehicle type observed by CAFE; and VMT_{ij} is the average annual vehicle-miles-traveled by model year and vehicle type in the I/M program area.

The formula normalizes predicted and observed emissions differences in I/M program and non-I/M program vehicles by model year to the on road fleet fraction and average annual mileage of that model year. This exercise enables the different units of measurement between on road and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form.

A.4 Analysis

This section reports the results of the reference method for evaluating the Atlanta enhanced I/M program during its third two years of operation. The evaluation uses remote sensing emissions data collected in 2010 and emission factors predicted for the 2010 fleet by an EPA computer model. The 2010 calendar year represents the end of the sixth full cycle of enhanced IM testing.

Because the reference method involves direct comparisons between on road data and EPA's MOBILE6.2 model, we restrict the data in several ways to obtain an "apples-to-apples" comparison. First, only 1986 to 2007 model year cars and trucks are included in the analysis. The 2008, 2009 and 2010 model years are not included since these vehicles were exempt from testing in 2010.

The second data restriction is the use of only vehicles registered in the thirteen Atlanta counties of the I/M program area as an *Inspected* fleet while some of the inspected vehicles could move to non I/M areas due to natural migration (such as change of ownership, etc.).

A.4.1 Data Overview:

The remote sensing data used by this evaluation were collected at twenty (20) Atlanta I/M program area sites and seven non-program area sites in Augusta and Macon.[34] Measurements in the I/M program area were conducted from January to December 2010, while the non-program area measurements were collected over 12 days in April, March, June, July, September and December.

AQG collected and identified 118,726 measurements from vehicles registered in thirteen-county area with a 2008 inspection. In the non-program areas, 29,171 measurements were collected from vehicles registered in the counties comprising Augusta and Macon and 8,192 measurements from 12 Atlanta counties that are not subject to the I/M program but are included in the Atlanta Clean Fuels program. The measurements in Augusta and Macon non-I/M areas are substantially greater than the 16,797 measurements collected in these areas as part of the 2002 evaluation[35]. This shift of measurements into the non-I/M areas was designed to improve the overall statistically validity of the reference

[34] I/M program area measurements are made within thirteen counties that comprise the metropolitan Atlanta ozone nonattainment area: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Fulton, Gwinnett, Henry, Forsyth, Paulding, and Rockdale. Non-I/M program measurements include Bibb, Richmond and Columbia counties.

[35] 2002 was the first year when Atlanta 13 counties were compared to two different control points. In following years number of measurements collected at control sites was increased to improve the quality of analysis.

comparison. In both program and non-program areas, we randomly selected one measurement from unique vehicles with multiple readings.

This evaluation of the Atlanta enhanced I/M program relies on measurements of CO and NOx data. The primary reason for focusing on CO over HC is that the former pollutants have a greater signal-to-noise ratio. CO's lower variance is due to its presence in higher concentrations than HC, making it easier to measure by remote sensing devices and less susceptible to weather and driving conditions. In other words, although HC data was collected and analyzed during the study, these data were not the primary focus of the analysis.

A.4.2 Validity of Fleet Comparisons:

Our ability to infer I/M effectiveness from the emission differences in Atlanta and Augusta/Macon vehicles hinges upon the comparability of the three fleets. The inspected Atlanta and uninspected Augusta-Macon vehicle fleets have similar model year distributions, although the inspected Atlanta fleet is slightly newer than the uninspected Augusta-Macon fleet (Figure A.1).

A second issue for the validity of our analysis is whether the Augusta and Macon fleets are similar enough to be combined into one uninspected fleet. Both this and the previous studies had shown that the average CO emissions by model year and vehicle type do not differ significantly between the two uninspected fleets and thus these results are combined.

The third issue is the ability of the Augusta/Macon fleet to serve as a reference for the estimation of I/M effectiveness in Atlanta. While this was a reasonable assumption for the first two evaluations, the introduction of low sulfur gasoline initiative in 25 Atlanta counties changed this. Vehicles registered in Atlanta 13-county area were thus able to seize benefits from two different emission control programs while vehicles from Augusta and Macon have neither. Thus for purposes of this evaluation, the Augusta-Macon fleet was used to analyze a combined effect of I/M and fuel programs and data collected from 12 Atlanta counties that are not subject to I/M testing was used for estimation of I/M contribution alone.

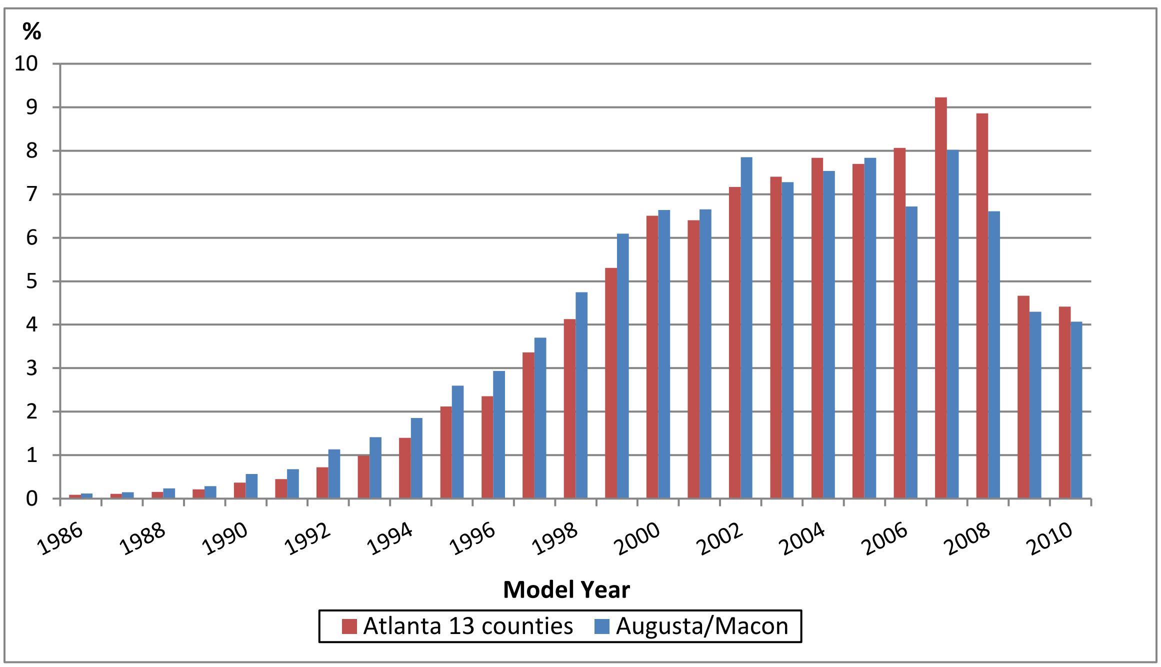

Figure A.1 Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.

A.4.3 Reference Method Results:

The results of the reference method for evaluating the effectiveness of the Atlanta enhanced I/M program are laid out in Tables A.1. But first, let us review the methodology for generating the estimates. We calculate the emissions difference in inspected and uninspected cars and trucks by model year and then weight those differences to that model year's annual average mileage and fleet fraction. The exercise is undertaken separately for predicted emissions factors and on road emissions data. The weighted emissions differences in each category are then summed over all model years. The weighted value based on onroad emissions data becomes the numerator, whereas the weight value based on predicted emission factors becomes the denominator. Dividing the numerator by the denominator yields the percentage of expected emissions differences actually achieved in inspected and uninspected vehicles. The results of this exercise indicate that the Atlanta enhanced I/M program captures 112 percent of CO reductions for cars and 72 percent for trucks compared to those predicted by EPA.

Table A. 1 Effectiveness of Atlanta I/M Program and Fuel Program.

	Atlanta 13-counties inspected fleet vs. Augusta-Macon uninspected fleet	
	Cars	**Trucks (LDT2)**
CO	112%	72%
NOx, lower estimate	117%	84%
NOx, higher estimate	135%	107%

Delving into the data comprising these results, Figure A.2 and Figure A.3 compare the CO emissions differences in inspected thirteen-county Atlanta and uninspected Augusta-Macon vehicles measured on road by RSD.

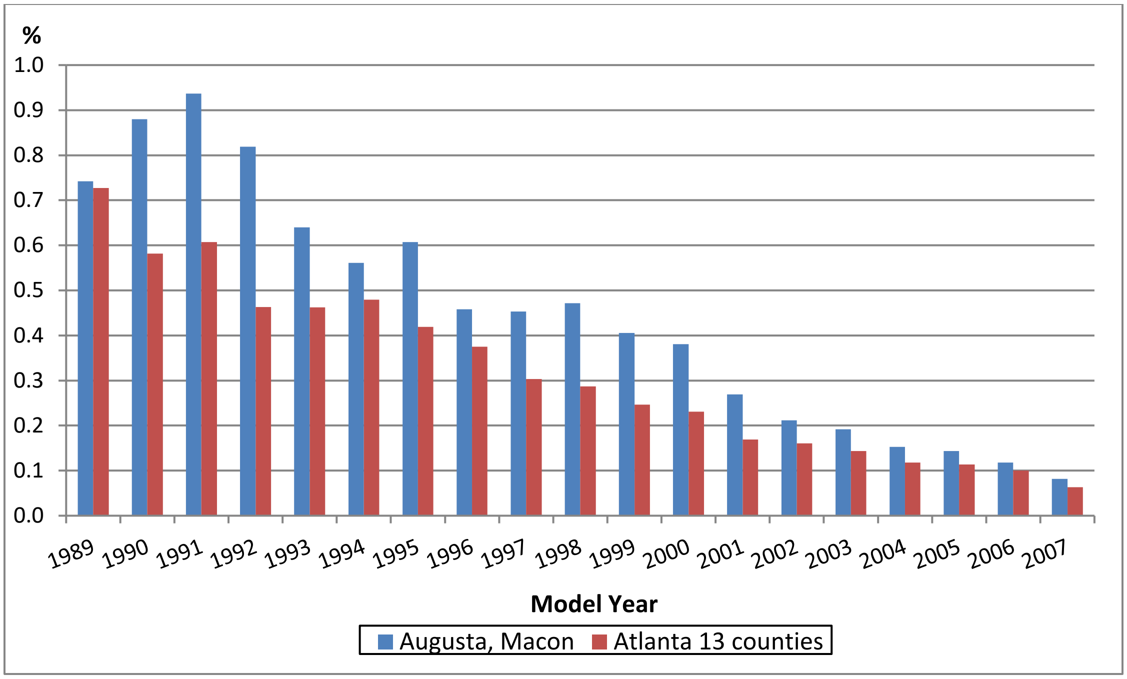

Figure A.2 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.

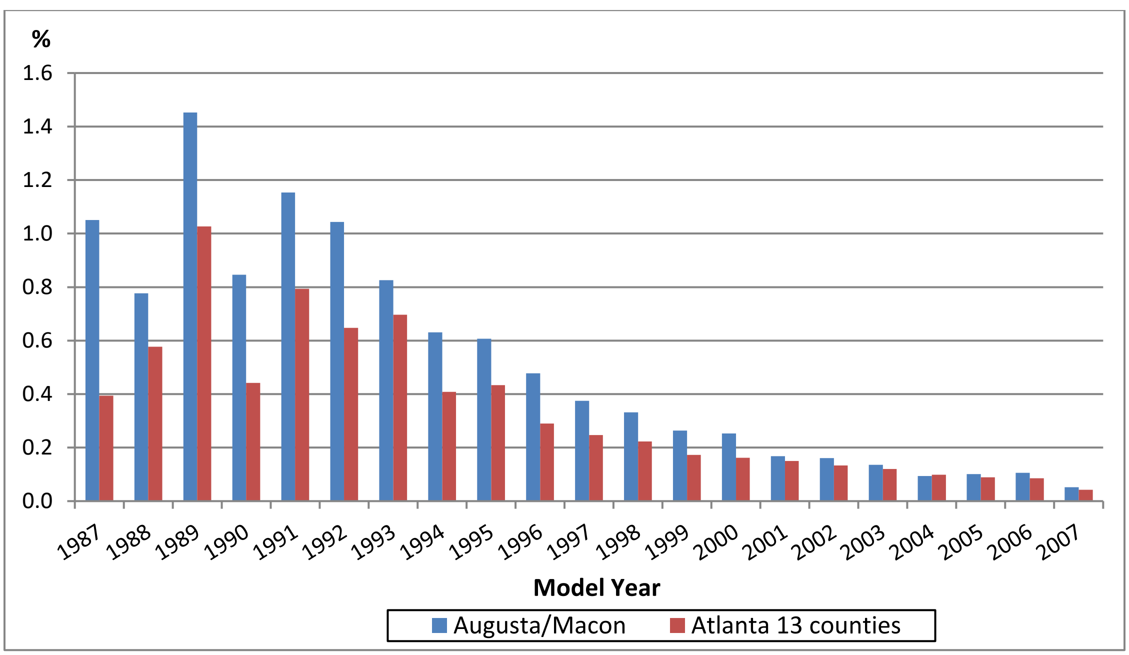

Figure A.3 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. LDT2 Trucks Only.

The on road emission differences for NOx mimic this pattern, although with much larger fluctuations due to additional benefits Atlanta gets from the usage of low sulfur fuel. It is known that the amount of sulfur in the gasoline affects level of NOx exhausted. Figure A.4 illustrates the changes in the average NOx values due to seasonal variations of sulfur level in the gasoline supplied. Therefore additional references are needed to separate I/M air quality benefits and those from low sulfur fuel.

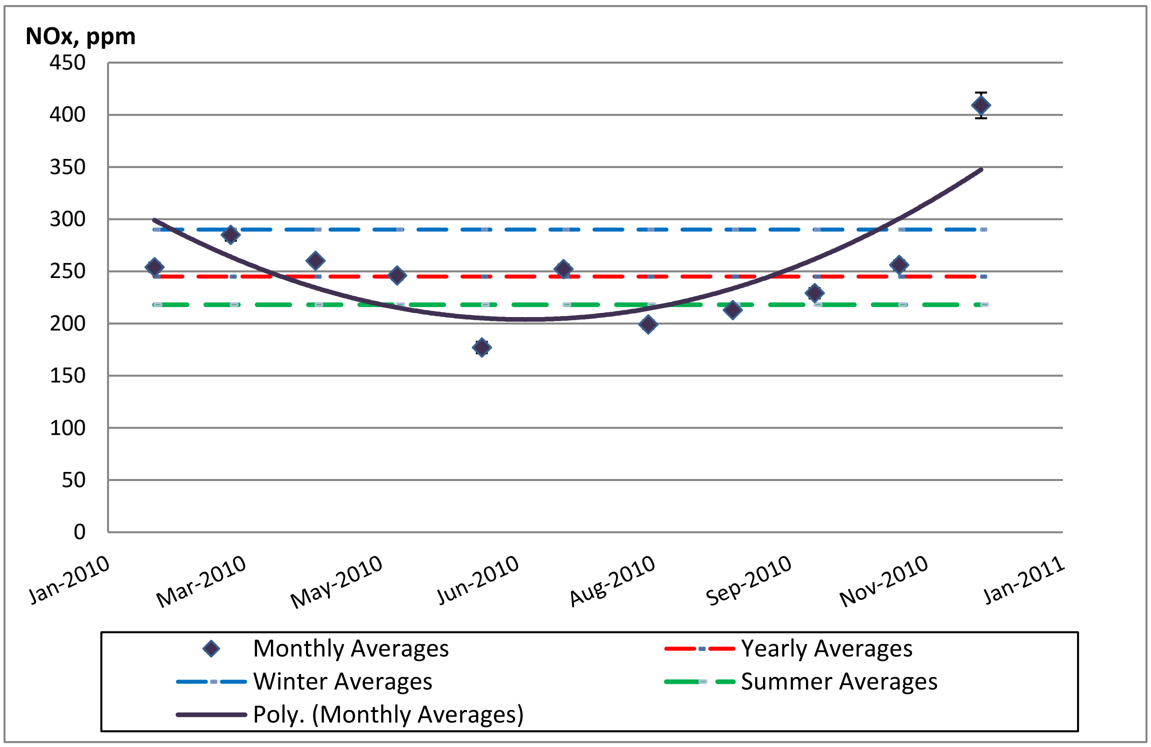

Figure A.4 Seasonal Average Values of NOx. All Passenger Vehicles (LDT & LDV). Atlanta 13 Counties.

The data collected from vehicles recorded in the twelve Atlanta MSA counties that are not subject to I/M but operate on the same fuel was used as one of such references and represents the lower point of our estimated I/M benefits for NOx. The usage of this Atlanta "uninspected" fleet as the solid mark for comparison does not seems justified. First, our evaluation of I/M records has shown that due to local migration between counties about 20% of the vehicles that are not subject to I/M in this area have actually undergone the testing during previous two years. Therefore comparing the Atlanta inspected and "uninspeced" fleets we will tend to underestimate the benefits from I/M program. On the other hand, it is perceived that these same 12 counties may have become a repository for the vehicles that are most likely to fail the I/M test. In other words, people who are trying to avoid I/M testing may be seeking to register their vehicles outside of the I/M area. However, our previous work indicates that number of such vehicles does not exceed 3% from the fleet. Based on the reasoning outlined above we determine the I/M effectiveness for NOx derived from direct comparison of Atlanta 13 counties and Atlanta 12 counties as the lower estimation point assuming that it represents only 80% of actual benefits and higher estimation results from adding an additional 20% benefit for the tested fraction.

A.5 Discussion

Interpreting emissions differences in the Atlanta inspected fleet and Augusta/Macon fleets as combined effect of the enhanced I/M program and fuel programs assumes that we have controlled for all differences in these fleets. This assumption is challenged by the possibility that the Augusta/Macon fleet is composed of higher mileage or poorer quality vehicles than the Atlanta thirteen-county fleet. One source of evidence for mileage differences, the U.S. Department of Transportation data on daily vehicle miles traveled (VMT), suggests that vehicles in Atlanta travel 34 miles per day per capita versus 22 miles for vehicles in Augusta. This information would seem to weaken any hypothesized mileage difference, at least between Augusta and Atlanta. However, because GDOT estimates are based on observed freeway traffic flows that capture out-of-state as well as local vehicles, it is difficult to extrapolate these VMT estimates to the local vehicle fleet. Exclusion of luxury cars from analyzed data sets did not make significant changes in emission patterns therefore the fleet composition differences between Atlanta and Augusta/Macon are negligible.

The comparison of Atlanta thirteen counties inspected fleet with Atlanta uninspected fleet has the same validity issues. Since vehicles that likely to fail testing have the tendency for migration into neighboring counties that are not subject for I/M program we may overestimate its effectiveness. But due to close proximity inspected vehicles also penetrate the noninspected area after change of ownership or under other circumstances which leads to underestimation of I/M benefits.

A.6 Comparing Results with Previous Reviews

The reference method for evaluating vehicle inspection/maintenance programs yields several advantages over other methods using on road remote sensing data. In fact, the reference method could be repeated over time to measure incremental effectiveness as more of the fleet is tested, inspectors become adept at identifying noncompliant vehicles, repair technicians gain experience at repairing emission control failures, and (more pessimistically) motorists learn better how to co-opt the test.

The study presented evaluates the seventh two-year period of the established in Atlanta thirteen counties I/M program. The first evaluation review covered the 1997-1998 years and the second evaluation covered the years 1999 and 2000. Both of these studies compared the Atlanta inspected fleet with an uninspected fleet in Augusta and Macon. The advent of the Atlanta Clean Fuels program required that the third evaluation, covering the years 2001 and 2002, incorporate a second reference area just outside of the thirteen county I/M area to account for these fuel differences. However, limited measurements in these areas during the third evaluation period resulted in evaluation uncertainties greater than desired and the CAFE program measurement program was modified to dramatically increase the number of vehicles measured in these areas, at the expense of reducing measurements in the I/M area. Table A.2 summarizes results from all six reviews.

As discussed earlier, the changes in the reference areas (and the incorporation of NOx measurements into the analysis) used in these evaluations makes direct comparisons between the first two (1998 and 2000) and the latter five (2002, 2004, 2006, 2008, 2010) evaluations difficult. However, the first group and latter groups may be compared with each other.

The reference method is not without its limitations, however. Selecting a comparable non-program fleet is a challenging task. While differences in fleet age and car/truck composition are relatively easy to account for between I/M and non-I/M fleets, discrepancies in maintenance trends, socioeconomic conditions and vehicle quality are difficult to discern. However, the emissions differences illustrated in Figure A.2 and Figure A.3 make a compelling case that both the I/M and clean fuels programs have a significant and positive impact on motor vehicle emissions in the Atlanta area.

Further, Table A.2 illustrates that these emissions differences are durable to the extent that similar results have been observed over an extended period. Table A.2 also reveals a significant difference between NOx benefits achieved for cars and for light duty trucks. While it is tempting to suggest that this is a systematic problem within the program, this may not be the case. The popularity of SUVs during the late 1990's and the early 2000's means that the 2010 - 2004 (and to a lesser extent the 2002) evaluations of truck emissions are dominated by vehicles for which low emissions reductions are expected. Likewise, for the same reason, the car estimates in 2010 represent, on average, an older fleet than for earlier evaluations. We may therefore hypothesize that differences between "car" and "truck" efficiencies may be an age effect in combination with, or instead of, an intrinsic difference in the effectiveness of the program for these two classes of vehicles. Establishing this relationship will be the subject for future studies.

Table A.2 Biennial I/M Effectiveness Estimated for 1998 - 2010 Measurement Years.

Estimated IM Effectiveness		**Cars**	**Light Trucks**
1998 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	87%	75%
	NOx	NA	NA
2000 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	84%	84%
	NOx	NA	NA
2002 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	166%	229%
	NOx Low estimation	78%	68%
	NOx High estimation	170%	150%

2004 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	256%	223%
	NOx Low estimation	142%	72%
	NOx High estimation	176%	90%
2006 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	146%	126%
	NOx Low estimation	165%	105%
	NOx High estimation	206%	131%
2008 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	134%	79%
	NOx Low estimation	101%	111%
	NOx High estimation	121%	133%
2010 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	112%	72%
	NOx Low estimation	117%	84%
	NOx High estimation	135%	107%

APPENDIX B. BIENNIAL EVALUATION OF THE EMISSIONS REDUCTION EFFECTIVENESS OF THE ATLANTA VEHICLE INSPECTION AND MAINTENANCE PROGRAM FOR 2007-2008

B.1 Introduction

B.1.1 Overview:

The 1990 Clean Air Act Amendments (CAAA) made sweeping changes in the scope and stringency of vehicle inspection/maintenance (I/M) programs. Driven by persistent growth in vehicle travel and chronic air pollution in the nation's largest metropolitan areas, the legislation requires "enhanced" I/M programs employ advanced testing technologies and procedures as a way to better ensure the operability of vehicle emission control system.[36] The law also requires biennial evaluation of enhanced I/M programs and on road measurement of inspected fleet emissions, but does not link together the two requirements (CAA Title I §182c3C; CAA Title I §182c3Bi; Title I§182c3Ci).

In the absence of an explicit legislative linkage, the National Research Council has recommended that I/M programs be evaluated using on road emissions data collected by remote sensing devices (RSD) (National Research Council, 2001). RSD uses infrared and ultraviolet technology to measure the emissions of in-use vehicles.[37] The NRC report cited several advantages of RSD data for I/M evaluation. First, RSD is a cost-effective source of evaluation data compared with the higher per-vehicle costs of advanced dynamometer

---

[36] Enhanced I/M programs were required in areas of the United States in serious, severe, or extreme nonattainment of federal ozone standards. Moderate nonattainment areas were required to implement the less rigorous basic I/M programs. Marginal nonattainment areas had no I/M requirement.

[37] Infrared technology is used to measure carbon monoxide and volatile organic compounds. Ultraviolet technology is used to measure nitrogen oxides

testing on a small sample of vehicles, the original evaluation approach recommended by federal regulators. RSD data can also capture trends that cannot be discerned through internal inspection records alone, such as motorists avoiding the program and pre-inspection maintenance behavior. RSD data can also be used for a variety of purposes in addition to I/M evaluation, including mobile source emission inventories, clean-screen programs that exempt low-emission vehicles from subsequent I/M testing, and high-emitter programs that target polluting vehicles for off-cycle inspection and repair.

In response to this growing interest, the U.S. Environmental Protection Agency (EPA) released draft guidance in July 2001 for the use of remote sensing data for I/M program evaluation (U.S. EPA, 2001). The document outlines equipment specifications and measurement procedures along with study design techniques and quality control measures. The document also discusses three methodologies for analyzing remote sensing data to determine I/M program effectiveness. *The comprehensive method* compares the onroad emissions of the vehicle fleet before and after scheduled I/M testing. *The step method* compares inspected with uninspected model year emissions during the first year of a new or upgraded I/M program. *The reference method* compares the emissions of the vehicle fleet located in an I/M area with that of a distantly located non-I/M area.

This paper employs the reference method to evaluate the enhanced I/M program of Atlanta, GA. This major metropolitan area in the southeastern United States is home to

thirteen counties in “serious” nonattainment of the federal ozone standard.[38] The Atlanta enhanced I/M program was implemented in October 1996 in this thirteen-county area, replacing a basic I/M program that had been operating in four of the thirteen counties since the early 1980s. We estimate the effectiveness of the new I/M program by comparing the RSD emissions of a sample of its inspected vehicles with that of a sample of vehicles registered in the Georgia cities of Augusta and Macon. The latter areas have demographics, climate and fleet characteristics similar to Atlanta, but do not operate an I/M program. The emissions difference in the inspected Atlanta and uninspected Augusta/Macon vehicle fleets are then compared with that predicted by the commonly used EPA MOBILE6.2 computer model. Viewing model-predicted emissions differences in inspected and uninspected vehicles as the Atlanta I/M program goal and observed onroad emission differences as actual program performance, we estimate I/M effectiveness as the ratio of these two numbers.

This section provides background on I/M programs, including an overview of I/M program operations, and a history of I/M programs in Atlanta. The second section reviews current enhanced I/M evaluation approaches (including the RSD methods outlined in recent EPA guidance) and their respective strengths and weaknesses. The third section describes

[38] The federal ozone standard is 0.12 ppm averaged on an hourly basis and 0.08 ppm averaged over an eight-hour basis.. Ozone concentration is one of the six National Ambient Air Quality Standards set by EPA to protect public health. (The remainder include carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, particulate matter-10, particular matter-2.5, and lead.) There are five levels of nonattainment for these pollutants, ranging from marginal to extreme, which are determined by the number of times air monitoring stations in an area detect pollutant levels above federal standards during a specific timeperiod.

reference method data sources and methodology. The fourth section reports the results of the reference method. The fifth section discusses the results, with particular attention paid to their caveats. The sixth section presents our conclusions about the reference method results for Atlanta and how it compares with previous I/M evaluations using remote sensing data and reference method.

B.1.2 Vehicle Inspection/Maintenance Programs:

Vehicle inspection/maintenance (I/M) programs seek, first and foremost, to ensure the effectiveness of vehicle emission control systems. The inspection process, which applies to light-duty vehicles of a certain age, involves scheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls.[39] Inspections can be provided by decentralized test-and-repair networks, which allow service stations and automotive repair shops to perform emissions tests and repair failed vehicles, or by centralized test-only networks, in which a limited number of centrally operated facilities perform testing as the sole service. Depending on program design, the test may be performed annually or biennially, while the vehicle idles or is placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of on road driving conditions.

[39] The model year vehicles subject to testing vary across I/M programs.

Motorists must repair failed vehicles, comprising the maintenance component of the program. Vehicles with repair costs above a set amount may qualify for a waiver -- an exemption from further repair and testing -- provided that attempted repairs show some emissions improvement and are not triggered by tampering. Compliance is typically verified through the presence of a vehicle windshield sticker received after passing the test or through the vehicle registration process that requires an emissions certificate.

B.1.3 Inspection/Maintenance In Atlanta:

Atlanta's first I/M program was established in 1981, covering the three ozone nonattainment area counties of Fulton, Cobb and DeKalb. The fast-growing Gwinnett County was added in 1986. The program was implemented through a decentralized test-and-repair network which allowed repair shops, service stations and automobile dealers to perform emission inspections and emissions-related repairs. Testing was originally required for the latest ten model year vehicles, but was expanded in 1986 to include the latest twelve model years. To receive an emissions compliance certificate, cars were required to pass an idle emissions test and an inspection of the catalyst, air pump and fuel inlet restrictor for evidence of tampering. Owners of failing cars that spent more than $50 for repairs qualified for a waiver and an emissions certificate, so long as repairs were not due to tampering and showed some emissions improvement. Owners of cars that failed the tampering inspection were required to obtain repairs to bring their emissions into compliance regardless of cost.

In response to the 1990 Clean Air Act Amendments (CAAA), the Georgia legislature revisited emissions testing in 1992[40]. This legislation enabled the Georgia Department of Natural Resources (GDNR) to upgrade Georgia's I/M program to an “enhanced” program, bringing it into compliance with the 1990 CAAA and new federal I/M federal regulations. This enhanced version of the program received limited implementation in October 1996[41], with emission inspections required only for those vehicles migrating to the Atlanta I/M program area. The new program commenced in January 1997, with biennial emissions testing required of all vehicles from the 1975 model year up until two years of age. The new program also spanned the 13-county nonattainment area, incorporating nine new counties that were not subject to the previous basic I/M program.

After the first two-years of operation several changes had been made to the program: the program began to require vehicles over six years of age to undergo the more rigorous Acceleration Simulation Mode (ASM) testing in October 1998, while in the first two-years period all vehicles were subjects of two-speed idle test (TSI)[42]. The primary difference between ASM and TSI testing is the approximation of real-world driving conditions, i.e., placing the engine under load. While the emissions inspector depresses the accelerator to

---

[40] 1992 Georgia Air Quality Act, Article 2: Motor Vehicle Emissions Inspection and Maintenance Act (OCGA Section 12-9-40 et seq.).

[41] October 1996 was chosen as the soonest possible start-up date after the previous basic I/M program, which operated during a January-to-April vehicle registration “season.” Vehicle registration is now conducted year-round in Georgia, as is enhanced emissions testing.

[42] two-speed idle (TSI) testing procedure that measures emissions under idle and a 2500 RPM engine speed

achieve 25 miles per hour (MPH), ASM testing places the vehicle's drive wheels on a treadmill-like dynamometer that applies a 25 percent load on the vehicle engine. The latter approach is more representative of actual driving conditions than an idle test. Vehicles that failed emissions testing were required to be brought into compliance by repair. Owners of covered vehicles in the 13-county ozone nonattainment area were required to show proof of a passing emissions inspection, a waiver, or proof that they qualify for an exemption in order to register their vehicle.

This report concentrates on the fifth two-year period of inspection and maintenance program operation in Atlanta and covers years 2005 and 2006. By this time certain significant changes have been made to the Atlanta enhanced I/M program. The waiver limit had been increased several times. In 2001, testing frequency changed from biennial to annual; the requirement to inspect vehicles back to 1975 model years was replaced with the requirement to inspect the latest 25 model years; and the exemption from testing of the newest two model years was changed to exemption of the newest three model years.

B.2 Enhanced I/M program Evaluation Methods

Three types of data currently dominate the evaluation of enhanced I/M programs: I/M records, which document the results of each inspection; roadside pullovers, which administer emissions tests to vehicles of randomly selected willing motorists; and remote sensing data, which measures on road vehicle emissions. This section reviews evaluations employing each data type, along with the strengths and weakness of each.

B.2.1 Emissions Inspection Records:

The most common source of biennial evaluation data comes from emissions test records generated by I/M programs.[43] I/M test records provide a cost-effective source of evaluation data because they are routinely generated and easily accessible. Because I/M records cover the entire inspected vehicle population, statistical conclusion validity is generally not an issue: evaluators can control for a variety of vehicle characteristics that influence emissions. The availability of odometer data in most I/M records is also advantageous, enabling evaluators to control for the influence of mileage on emissions. A final advantage stems from inspection/maintenance protocols, which are designed to correlate with the Federal Test Procedure[44] and to facilitate quality control.

However, I/M records suffer from weaknesses that limit their reliability as the sole indicator of program performance. Chief among these is the inability to parcel out fraudulent testing behavior, particularly when inspectors substitute clean-emitting vehicles for unrepaired high-emitting ones on the retest (Wenzel et al., 2000). I/M records may also underestimate program effectiveness by missing pre-inspection maintenance performed by some motorists to lower I/M test failure risks. While it is difficult to quantify the impact of

[43] Personal conversation with James Lindner of the U.S. Environmental Protection Agency's Office of Transportation and Air Quality, September 26, 2001. Also see http://www.epa.gov/otaq/epg/progeval.htm.

[44] The Federal Test Procedure is an elaborate testing protocol established in the early 1970s to certify manufacturer compliance with the 1970 Clean Air Act-mandated new vehicle emission standards.

such maintenance, it is expected to yield artificially low baseline emissions and thus underestimate program effectiveness. Generally speaking, these weaknesses speak to the role of I/M records as an internal, not an independent, source of evaluation data.

Evaluations employing I/M records also make tradeoffs between internal validity and representativeness of the data. The inspection process employs highly-controlled conditions to ensure that vehicles are measured under consistent circumstances (e.g., engine stress, vehicle speed, and temperature). While these controls reduce confounding influences on emissions, they represent only a fraction of driving conditions that typify onroad driving. Consequently, the ability to extrapolate I/M test emissions to onroad emissions is limited.

To estimate I/M effectiveness, some evaluations calculate the average emissions difference between the initial and final test scores on failing vehicles and assume that the difference is attributable to the I/M program. Three studies used this approach to evaluate different time periods of the Arizona enhanced I/M program. Two of these studies (Wenzel, 1999 and Glover and Brzezinksi, 1999) estimated a 14 percent reduction in carbon monoxide (CO), a 15 percent reduction in hydrocarbons (HC), and a seven percent reduction in nitrogen oxides (NOx). The third study (Ando, et al, 1999), focusing on repaired vehicles, estimated emission reductions of eight, eight and fourteen percent for CO, HC and NOx.

Sierra Research (1998) also compared initial and final emission results for failed vehicles in AirCare, the Canadian Vancouver/British Columbia emissions testing program.

This study estimated I/M emission reductions of 13 percent CO, 9 percent HC, and 4 percent NOx. Replacing initial emission results of failed inspections with EPA model predictions of an untested fleet's emissions, the researchers estimated 16 percent, 20 percent, and 14 percent emission reductions for CO, HC and NOx. The latter emission differences are thought to be higher than the former because model predictions, as opposed to initial inspection results, are not influenced by pre-inspection maintenance behavior.

The Colorado enhanced I/M program was twice evaluated using inspection records. The first analysis, comparing final test scores for vehicles inspected in 1997 with the new program's first 2,138 initial inspection test scores in 1995, indicated CO emission reductions in the range of 30 to 34 percent (Environ, 1998). The second analysis compared failed vehicles' initial and final inspection results from 1998 that had been converted to Federal Test Procedure scores. The comparison, which normalized repair benefits to the entire inspected fleet, suggested that CO had been reduced by eight percent and HC by six percent, with NOx increasing by one percent (Office of the State Auditor, 1999). While the study results seem contradictory, they cover different timeframes, make divergent assumptions (about deterioration rates, the fate of vehicles with final failures) and employ different measures in estimating I/M effectiveness.

One weakness in attributing before-after emission differences to I/M is the potential for "regression to the mean" emissions behavior, in which a portion of I/M failures will

register lower emissions on the final inspection without repair.[45] This phenomenon is driven by tremendous emissions test-to-test variability, the presence of vehicles with marginally failing emissions, and variance in environmental conditions favorable to test performance. Without verifying repairs, the emissions differences between initial and final test scores may overestimate program effectiveness.

B.2.2 Roadside Emission Inspections:

Used primarily in California, roadside emissions tests are administered with the aid of law enforcement officers who randomly pull vehicles over and ask motorists to voluntarily submit their vehicles to an emissions inspection. Volunteer license plate numbers are then used to query the I/M program database to determine those vehicles with and without an inspection in the past twelve months. Recently inspected and uninspected vehicle emissions are then compared to estimate the emission reductions due to enhanced I/M. Roadside emissions estimates of 1999 enhanced I/M program effectiveness indicate emission reductions of 13 percent for CO, 14 percent for HC, and 6 percent for NO (California Air Resources Board, 2000).

[45] Regression to the mean occurs when two imperfectly correlated measures are compared for a nonrandom sample. The nonrandom sample is typically drawn from high or low scorers on either measure. Regression to the mean occurs when the sample mean moves towards the population mean in the absence of intervention. In the context of I/M evaluation, this means that certain vehicles failing their initial I/M test will score more closely to the mean of the population on the retest, i.e., register passing emissions, without repair. Regression-to-the mean can also occur in vehicles that pass their initial inspection but would fail a subsequent retest.

As with I/M program data, roadside pullovers enable the collection of odometer data for mileage estimates. In contrast with I/M program data, the spontaneity of roadside inspections preclude fraudulent test results that overestimate effectiveness, as well as pre-inspection maintenance behavior that underestimates program effectiveness. However, because roadside emissions tests employ a portable version of official inspection procedures, they sacrifice real-world driving conditions. Furthermore, the approach is costly and generates limited data, requiring as many as four technicians and one law enforcement officer to measure approximately 25 vehicles per day (Wenzel, et al 2000, p. III-8). Self-selection bias is a risk because the test is voluntary and tends to yield a ten percent refusal rate (Wenzel, et al 2000, p. III-8).[46]

B.2.3 Remote Sensing Data from Onroad Vehicles:

A second source of data for evaluating I/M program effectiveness, the one used in this study, is from remote sensing devices (RSD) that measure the emissions of vehicles while they are being driven. The advantage of in-transit measurement is the ability to observe a vehicle's emissions under typical driving conditions, which cannot be as easily captured by traditional controlled emissions testing procedures. Remote sensors can measure a large number of vehicles, an important attribute given the need to control for

[46] The evidence of such bias is mixed. One recent study that used remote sensing to measure the vehicle emissions of refusals and participants alike found no significant difference between the two groups (Wenzel et al, 2000, pg. III-8), while an earlier similar study found that refusal vehicles had 2.5 times the emissions of volunteer vehicles (Stedman, 1994).

tremendous emissions variability due to vehicle type, age, make and model, and emission control technology. A final advantage stems from the unscheduled nature of the measurement, which precludes pre-inspection and fraudulent maintenance behavior that can occur when motorists (as with I/M tests) know when a measurement will occur.

In contrast with the highly controlled parameters of the emissions inspection, the physical circumstances of remote sensing data collection are only approximated through sampling site characteristics (e.g., moderate grades to ensure vehicles operate under only a slight engine load and sampling sites that avoid residential areas to minimize inflated emissions from cold engines). Another drawback is that remote sensors capture a split-second emissions reading that may not reflect a vehicle's typical emissions, making larger samples sizes a requirement to average out random emission fluctuations and to profile emissions aggregated within vehicle type (cars vs. trucks) and model year.

Remote sensing data has been used in three ways to evaluate I/M programs. The first method averages the emissions of vehicles measured before initial and after final I/M testing, with the difference attributed to I/M program effectiveness. Dubbed the "comprehensive method" in recent EPA evaluation guidance, emissions differences can also be generated for various subfleets, such as vehicles initially failing and ultimately passing I/M testing versus failing vehicles that never receiving a final pass. This approach enables a variety of I/M-related analyses, such as deterioration rates of I/M repairs, the influence of pre-I/M repairs on emissions baselines, and a comparison with estimates based on I/M records alone. The major disadvantage to this approach is the enormous volume of onroad data required to measure a representative sample of vehicles before and after I/M

testing. Sample size requirements hinge on the probability of measuring vehicles onroad within a specific timeperiod of I/M testing, a probability that fluctuates with testing frequency and the distribution of sampling throughout the year.

The comprehensive method was used to estimate the effectiveness of the California South Coast Air Basin's enhanced I/M program in 1999 (Wenzel et al., 2000). "Smog Check" I/M records were used to delineate tested from untested vehicles by the existence of an enhanced inspection within the past twelve months.[47] A comparison of these vehicle groups indicates a ten percent reduction in CO, a four percent reduction in HC, and a five percent increase in NOx. An earlier remote sensing study in California in 1996 compared the onroad emissions of 3.5 million vehicles 30 to 90 days before with up to 90 days after their basic I/M test (Klausmeier and Weyn, 1997). For those vehicles that failed their initial smog check and then passed, both CO and HC emission differences registered at 20 percent. Normalizing this result to the entire fleet yielded an estimated nine percent emissions reduction in HC and CO. A third evaluation, of the Arizona enhanced I/M program in 1997, analyzed four million remote sensing measurements on 1.2 million vehicles in the Phoenix I/M area (Wenzel, 1999). The results indicated a seven percent reduction in CO and an 11 percent reduction in HC.

[47] Untested vehicles may have been inspected under the previous basic I/M program more than twelve months ago or they may have had an enhanced inspection after the remote sensing reading.

One weakness in the comprehensive method is the potential seasonal effects that results from the year-round testing required to obtain adequately sized samples. Users of this method have also tended to rely on a few high-volume sites, yielding a large number of repeat vehicles that lower the fraction of unique vehicles that could be reached at a greater number of sites.

A second I/M evaluation approach using remote sensing, known as the Step Method, compares inspected with uninspected vehicles during the first year of a new or upgraded program. The uninspected vehicles comprise an internal control group against which to compare the emission reductions of the inspected vehicles. Because the method applies to the early phases of a new or improved program, it can be used only once to assess program effectiveness.

A remote sensing study of the Colorado Enhanced I/M program compared odd (inspected) and even (uninspected) model year vehicles during the end of the first year of a new biennial enhanced I/M program (Stedman, et al, 1997). At that point, in program history, all odd model year vehicles should have been inspected, whereas all even model year vehicles had no reason to be inspected. This timing rendered even model year vehicles the untested control group against which to compare the odd model year vehicle emissions. The comparison of odd and even model year emissions suggested that Colorado's enhanced I/M program had reduced CO between five and nine percent, while HC and NO showed no improvement.

Three factors limit the generalizability of the Colorado study results to enhanced I/M program effectiveness. Remote sensing took place in a single location, which avoids any confounding socioeconomic or physical influences at different sites but limits generalizability to the overall fleet. Furthermore, vehicles traveling past the remote sensing site were decelerating, which does not represent typical driving conditions and is not the optimal condition for measuring carbon monoxide (Environ, 1998, p. 2-19). A third limitation was that the study measured vehicles transitioning from an annual basic I/M program to an enhanced I/M program, rendering it an evaluation of incremental program effectiveness and not a complete estimate of enhanced I/M program performance.

A third approach using remote sensing data (the one used in this study) compares the onroad emissions of vehicles registered in an I/M area to that of vehicles registered in non-I/M areas. The non-I/M area serves as a surrogate untested fleet. The validity of this approach relies on the selection of a non-I/M area comparable in fleet age, a well-documented contributor to vehicle emissions; climate, which can accelerate emission control equipment deterioration; and demographics, which influences the age, quality, and maintenance of the vehicle fleet. This approach was originally applied to the basic I/M program operating in four counties of the thirteen-county Atlanta ozone nonattainment area, with the nine nonattainment counties without I/M comprising the untested fleet. The analysis indicates that car and truck emissions for CO were 15 and ten percent higher, respectively, in the uninspected nine-county fleet than in the inspected four-county basic I/M fleet. The study is limited by its inability to control for differences in mileage and socioeconomic conditions between the two vehicle fleets.

B.3 I/M Program Evaluation Components

This study employs I/M program evaluation method that compares the on road emissions differences *observed* in inspected and uninspected vehicles with the same emissions differences *predicted* by a U.S. Environmental Protection Agency mobile emissions model. The model-predicted emissions difference represents the goal of the I/M program, a reasonable assumption given that states use the model to generate the emission reduction credit received for automobile emissions testing programs. The emissions difference observed in on road inspected and uninspected vehicles is assumed to reflect I/M program performance, an assumption rendered plausible only by the comparability of the inspected and uninspected fleets. We will devote attention in the next section to answering the comparability question.

This section describes the collection of data used in the evaluation. It details the Continuous Atlanta Fleet Evaluation (CAFE), the remote sensing study of on road Georgia vehicles that provides on road emissions data of inspected and uninspected vehicles. The MOBILE6, EPA's recommended emissions model, from which we extracted predicted emission factors, is also discussed. The last section outlines the algorithm that combines data from CAFE and MOBILE6 to generate effectiveness estimates for the Atlanta enhanced I/M program.

B.3.1 On Road Emissions Data:

The Continuous Atlanta Fleet Evaluation (CAFE) provides the on road emissions data used to represent, *inter alia*, Atlanta enhanced I/M program performance. CAFE uses

remote sensing devices to measure annually the emissions of approximately 380,000 in-use vehicles in the 13-county I/M program area, as well as two cities located more than 75 miles from Atlanta that do not require vehicle emissions testing.[48] The study is an ongoing effort started in 1993 to collect vehicle emissions data for assessing a variety of trends, including fleet turnover, emission control deterioration, and socioeconomic impacts of mobile source control strategies.

RSD measures the emissions of passing vehicles remotely and unobtrusively so motorists are minimally aware of the equipment and do not alter their natural driving behavior. To that end, the remote sensing instrumentation is housed in a van parked on the roadside along with a video camera. An infrared light source and its generator are placed on the opposite side of the road or on the median to create a beam of light that traverses the road. When a passing vehicle breaks the beam, it triggers a measurement of hydrocarbons, carbon monoxide, and nitrogen oxides in the exhaust. Simultaneously, a video camera records the vehicle's license plate, which is automatically scanned into the database of emissions measurements.

After data collection, remote sensing measurements are merged with vehicle registration records using the vehicle license plate. The resulting database allows various characteristics of measured vehicles to be identified, including vehicle identification

[48] Augusta is located 136 miles east of Atlanta, whereas Macon is 76 miles south of Atlanta.

number,[49] make, model year, and vehicle type. License plates are also linked with inspection/maintenance records to identify vehicles with prior emission inspections.

RSD sampling sites are selected to ensure physically consistent but demographically diverse characteristics. Single straight lines of traffic with an average 35 mile-per-hour velocity are sought to facilitate single vehicle measurements and speeds that maximize measurement opportunities. Driver behavior and driving maneuvers are also observed at each site to ensure that remote sensing measurements would not be biased high by acceleration or low by coasting. Finally, notations are made during the site visits regarding any obvious or suspected diurnal patterns that exist which affect the traffic volume. If distinct variations are found to exist in sites ultimately selected, sampling times are scheduled to account for those diurnal patterns. U.S. Census tract data and traffic count reports inform the selection of different income ranges and land uses.

The remote sensing sites relevant to this study reside within the 13-county Atlanta I/M program area, 12 Atlanta counties without an I/M program but subject to the Atlanta clean fuel program, as well as the Georgia cities of Augusta and Macon that have neither program. The latter locales do not require emissions testing and thus provided an uninspected vehicle fleet to serve as a control group for our previous I/M evaluations.

[49] Vehicle identification numbers are 17-digit alphanumeric strings that uniquely identify every vehicle manufactured. When decoded, they provide additional characteristics on vehicles. The VIN-decoded data of particular relevance to this research are vehicle type (car, truck, multi-purpose vehicle, van) and model year.

These cities were chosen after a review of census data and registration records revealed them to have characteristics – median household income, population density, and fleet distribution -- most similar to Atlanta than three other Georgia cities considered. But for the reasons which will be explained later for the present analysis we also used as the reference point the data collected from the vehicles registered in 12 counties that surround Atlanta I/M program area.

According to the state regulation, effective April 1, 1999 the sulfur content of all gasoline supplied in a 25 Atlanta region[50] shall not exceed a seasonal average of 150 ppm (by weight) and, effective April 1, 2001, a per-gallon cap of 500 ppm (by weight)[51]. This rule made vehicle operational conditions in Atlanta 13-county nonattainment area and Augusta-Macon significantly unequal. Since there is no mechanism to separate benefits received from the usage of low sulfur gasoline and emission reductions due to I/M program, the usage of Augusta-Macon fleet as a control group for I/M program evaluation became questionable. In the effort to eliminate the fuel effect, the data collected from the vehicles registered in twelve counties[52] that are not subject to the I/M program but receive the same fuel as Atlanta 13-counties I/M program area have been used in the present analysis as a

[50] 25-county Atlanta region include 13-county I/M program area and 12 additional counties without I/M program: Barrow, Bartow, Butts, Carroll, Dawson, Hall, Haralson, Jackson, Newton, Pickens, Spalding, Walton.

[51] Rules for Air Quality Control Chapter 391-3-1, July 20, 2005. http://www.gaepd.org/Files_PDF/rules/rules_exist/391-3-1.pdf

[52] Eight of these twelve counties are in the Atlanta Metropolitan Statistical Area and thus are considered “Atlanta-area” counties.

reference point. The data collected on Augusta-Macon sites represents combined benefits from both I/M and GA fuel programs.

B.3.2 Predicted Emission Factors:

We used MOBILE 6.2, EPA's recommended computer model for estimation of mobile emission factors, to predict emissions differences in inspected and uninspected vehicles.

B.3.3 Evaluation Algorithm:

We estimated Atlanta enhanced I/M program effectiveness by comparing EPA model-predicted emission differences with observed emission differences in inspected and uninspected vehicles. The comparison yields a percentage that represents the proportion of expected emission reductions actually achieved by the program. The formula for estimating I/M effectiveness is as follows:

$$\text{Effectiveness}=\frac{\sum_{ij}[(O_{n_{ij}}-O_{m_{ij}})/O_{n_{ij}}](P_{n_{ij}})(C_{ij})(VMT_{ij})}{\sum_{ij}(P_{n_{ij}}-P_{m_{ij}})(C_{ij})(VMT_{ij})}$$

where: O_m and O_n are the average onroad emissions observed for a particular model year and vehicle type for I/M program and non-program vehicles, respectively; P_m and P_n are the model-estimated emission factors for I/M program and non-program vehicles for a given model year a vehicle type; C_{ij} is the fraction of the Atlanta fleet of that model year

and vehicle type observed by CAFE; and VMT_{ij} is the average annual vehicle-miles-traveled by model year and vehicle type in the I/M program area.

The formula normalizes predicted and observed emissions differences in I/M program and non-I/M program vehicles by model year to the on road fleet fraction and average annual mileage of that model year. This exercise enables the different units of measurement between on road and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form.

B.4 Analysis

This section reports the results of the reference method for evaluating the Atlanta enhanced I/M program during its third two years of operation. The evaluation uses remote sensing emissions data collected in 2008 and emission factors predicted for the 2008 fleet by an EPA computer model. The 2008 calendar year represents the end of the sixth full cycle of enhanced IM testing.

Because the reference method involves direct comparisons between on road data and EPA's MOBILE6.2 model, we restrict the data in several ways to obtain an "apples-to-apples" comparison. First, only 1984 to 2005 model year cars and trucks are included in the analysis. The 2006, 2007 and 2008 model years are not included since these vehicles were exempt from testing in 2008.

The second data restriction is the use of only vehicles registered in the thirteen Atlanta counties of the I/M program area as an Inspected fleet while some of the inspected vehicles could move to non I/M areas due to natural migration (such as change of ownership, etc.).

B.4.1 Data Overview:

The remote sensing data used by this evaluation were collected at twenty (20) Atlanta I/M program area sites and seven non-program area sites in Augusta and Macon.[53] Measurements in the I/M program area were conducted from January to December 2008, while the non-program area measurements were collected over 12 days in April, March, June, July, September and December.

AQG collected and identified 127,830 measurements from vehicles registered in thirteen-county area with a 2008 inspection. In the non-program areas, 23,971 measurements were collected from vehicles registered in the counties comprising Augusta and Macon and 10,521 measurements from 12 Atlanta counties that are not subject to the I/M program but are included in the Atlanta Clean Fuels program. The measurements in both of these non-I/M areas are substantially greater than the 16,797 and 8,398 measurements collected in these areas respectively as part of the 2002 evaluation[54]. This

[53] I/M program area measurements are made within thirteen counties that comprise the metropolitan Atlanta ozone nonattainment area: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Fulton, Gwinnett, Henry, Forsyth, Paulding, and Rockdale. Non-I/M program measurements include Bibb, Richmond and Columbia counties.

[54] 2002 was the first year when Atlanta 13 counties were compared to two different control points. In following years number of measurements collected at control sites was increased to improve the quality of analysis.

shift of measurements into the non-I/M areas was designed to improve the overall statistically validity of the reference comparison. In both program and non-program areas, we randomly selected one measurement from unique vehicles with multiple readings.

This evaluation of the Atlanta enhanced I/M program relies on measurements of CO and NOx data. The primary reason for focusing on CO over HC is that the former pollutants have a greater signal-to-noise ratio. CO’s lower variance is due to its presence in higher concentrations than HC, making it easier to measure by remote sensing devices and less susceptible to weather and driving conditions. In other words, although HC data was collected and analyzed during the study, these data were not the primary focus of the analysis.

B.4.2 Validity of Fleet Comparisons:

Our ability to infer I/M effectiveness from the emission differences in Atlanta and Augusta/Macon vehicles hinges upon the comparability of the three fleets. The inspected Atlanta and uninspected Augusta-Macon vehicle fleets have similar model year distributions, although the inspected Atlanta fleet is slightly newer than the uninspected Augusta-Macon fleet (Figure B.1).

A second issue for the validity of our analysis is whether the Augusta and Macon fleets are similar enough to be combined into one uninspected fleet. Both this and the previous studies had shown that the average CO emissions by model year and vehicle type do not differ significantly between the two uninspected fleets and thus these results are combined.

The third issue is the ability of the Augusta/Macon fleet to serve as a reference for the estimation of I/M effectiveness in Atlanta. While this was a reasonable assumption for the first two evaluations, the introduction of low sulfur gasoline initiative in 25 Atlanta counties changed this. Vehicles registered in Atlanta 13-county area were thus able to seize benefits from two different emission control programs while vehicles from Augusta and Macon have neither. Thus for purposes of this evaluation, the Augusta-Macon fleet was used to analyze a combined effect of I/M and fuel programs and data collected from 12 Atlanta counties that are not subject to I/M testing was used for estimation of I/M contribution alone.

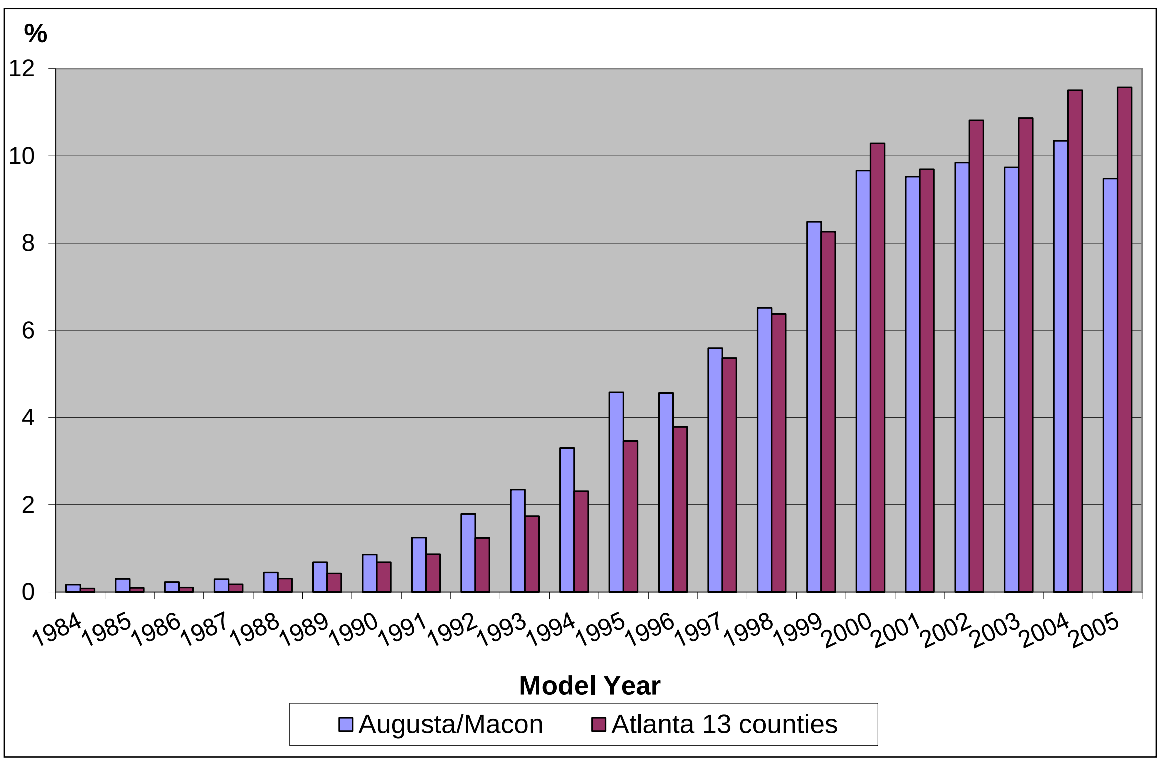

Figure B.1 Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.

B.4.3 Reference Method Results:

The results of the reference method for evaluating the effectiveness of the Atlanta enhanced I/M program are laid out in Table B.1. But first, let us review the methodology for generating the estimates. We calculate the emissions difference in inspected and uninspected cars and trucks by model year and then weight those differences to that model year's annual average mileage and fleet fraction. The exercise is undertaken separately for predicted emissions factors and on road emissions data. The weighted emissions differences in each category are then summed over all model years. The weighted value based on onroad emissions data becomes the numerator, whereas the weight value based on predicted emission factors becomes the denominator. Dividing the numerator by the denominator yields the percentage of expected emissions differences actually achieved in inspected and uninspected vehicles. The results of this exercise indicate that the Atlanta enhanced I/M program captures 134 percent of CO reductions for cars and 79 percent for trucks compared to those predicted by EPA.

Table B.1 Effectiveness of Atlanta I/M Program and Fuel Program.

	Atlanta 13-counties inspected fleet vs. Augusta-Macon uninspected fleet	
	Cars	**Trucks (LDT2)**
CO	134%	79%
NOx, lower estimate	101%	111%
NOx, higher estimate	121%	133%

Delving into the data comprising these results, Figures B.2 and B.3 compare the CO emissions differences in inspected thirteen-county Atlanta and uninspected Augusta-Macon vehicles measured on road by RSD.

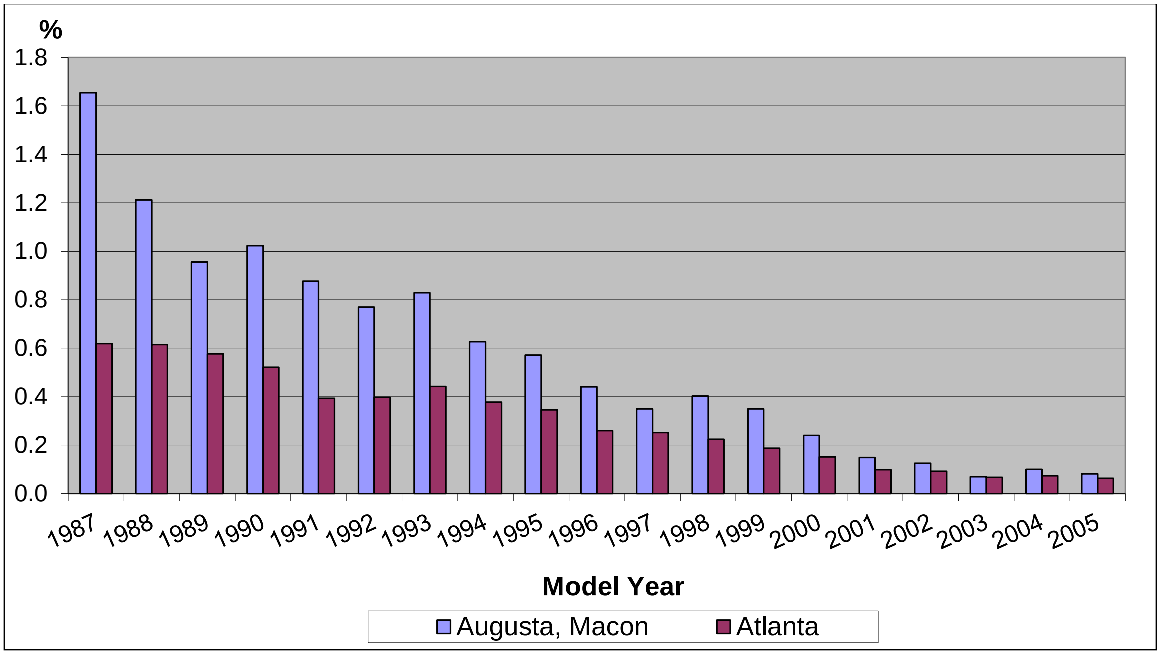

Figure B.2 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.

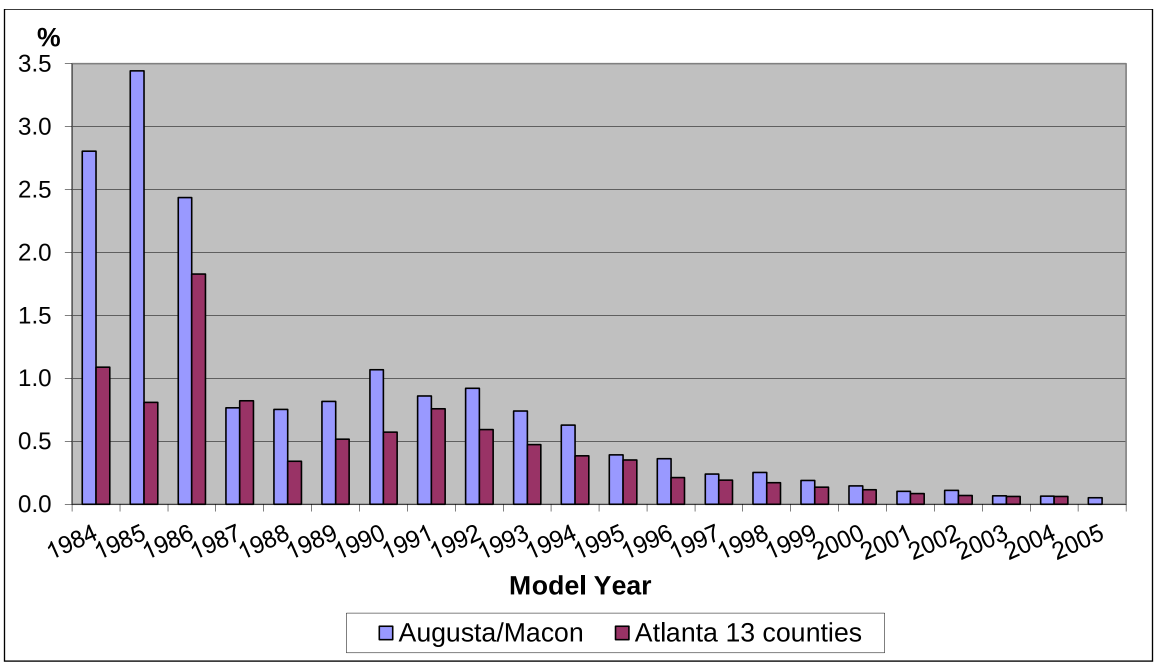

Figure B.3 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. LDT2 Trucks Only.

The on road emission differences for NOx mimic this pattern, although with much larger fluctuations due to additional benefits Atlanta gets from the usage of low sulfur fuel. It is known that the amount of sulfur in the gasoline affects level of NOx exhausted. Figure B.4 illustrates the changes in the average NOx values due to seasonal variations of sulfur level in the gasoline supplied. Therefore additional references are needed to separate I/M air quality benefits and those from low sulfur fuel.

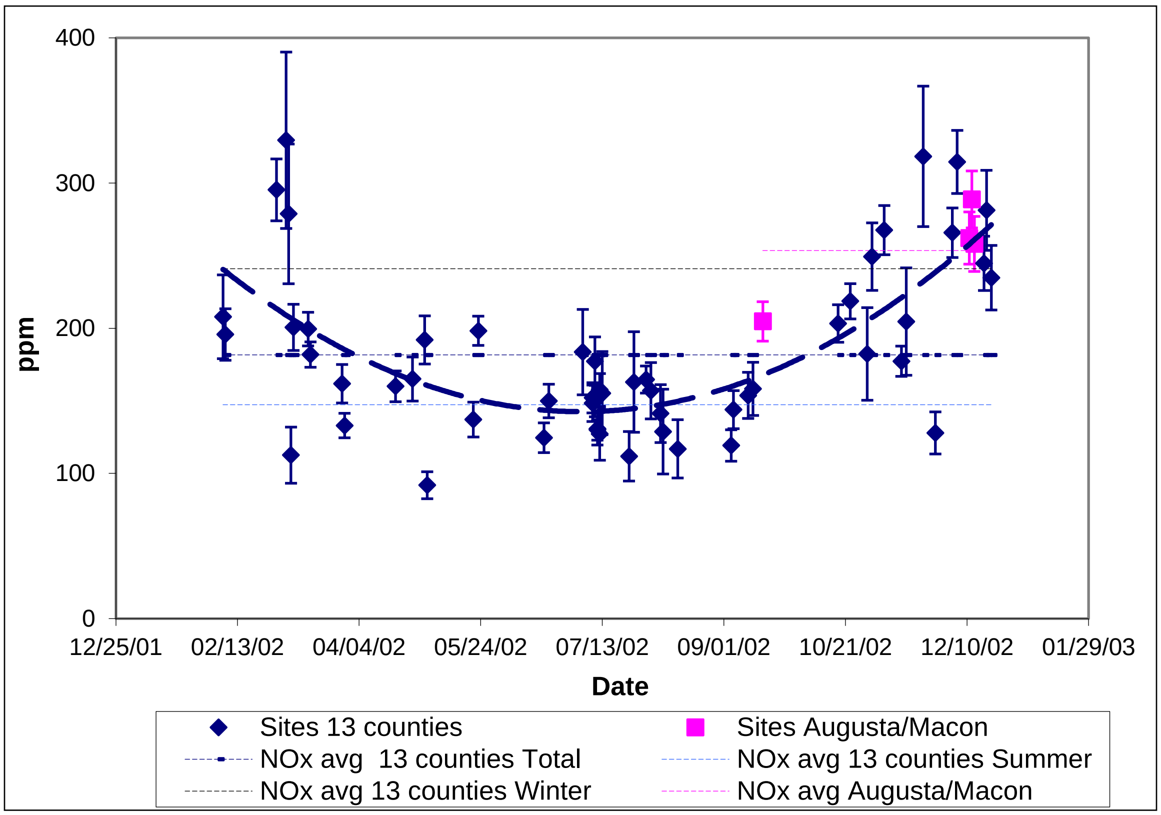

Figure B.4 Seasonal Average Values of NOx. Passenger Cars Only.

The data collected from vehicles recorded in the twelve Atlanta MSA counties that are not subject to I/M but operate on the same fuel was used as one of such references and represents the lower point of our estimated I/M benefits for NOx. The usage of this Atlanta “uninspected” fleet as the solid mark for comparison does not seems justified. First, our evaluation of I/M records has shown that due to local migration between counties about 20% of the vehicles that are not subject to I/M in this area have actually undergone the testing during previous two years. Therefore comparing the Atlanta inspected and “uninspeced” fleets we will tend to underestimate the benefits from I/M program. On the other hand, it is perceived that these same 12 counties may have become a repository for the vehicles that are most likely to fail the I/M test. In other words, people who are trying to avoid I/M testing may be seeking to register their vehicles outside of the I/M area. However, our previous work indicates that number of such vehicles does not exceed 3% from the fleet. Based on the reasoning outlined above we determine the I/M effectiveness for NOx derived from direct comparison of Atlanta 13 counties and Atlanta 12 counties as the lower estimation point assuming that it represents only 80% of actual benefits and higher estimation results from adding an additional 20% benefit for the tested fraction.

B.5 Discussion

Interpreting emissions differences in the Atlanta inspected fleet and Augusta/Macon fleets as combined effect of the enhanced I/M program and fuel programs assumes that we have controlled for all differences in these fleets. This assumption is challenged by the possibility that the Augusta/Macon fleet is composed of higher mileage or poorer quality vehicles than the Atlanta thirteen-county fleet. One source of evidence for mileage differences, the U.S. Department of Transportation data on daily vehicle miles traveled (VMT), suggests that vehicles in Atlanta travel 34 miles per day per capita versus 22 miles for vehicles in Augusta. This information would seem to weaken any hypothesized mileage difference, at least between Augusta and Atlanta. However, because GDOT estimates are based on observed freeway traffic flows that capture out-of-state as well as local vehicles, it is difficult to extrapolate these VMT estimates to the local vehicle fleet. Exclusion of luxury cars from analyzed data sets did not make significant changes in emission patterns therefore the fleet composition differences between Atlanta and Augusta/Macon are negligible.

The comparison of Atlanta thirteen counties inspected fleet with Atlanta uninspected fleet has the same validity issues. Since vehicles that likely to fail testing have the tendency for migration into neighboring counties that are not subject for I/M program we may overestimate its effectiveness. But due to close proximity inspected vehicles also penetrate the noninspected area after change of ownership or under other circumstances which leads to underestimation of I/M benefits.

B.6 Comparing Results with Previous Reviews

The reference method for evaluating vehicle inspection/maintenance programs yields several advantages over other methods using on road remote sensing data. In fact, the reference method could be repeated over time to measure incremental effectiveness as more of the fleet is tested, inspectors become adept at identifying noncompliant vehicles, repair technicians gain experience at repairing emission control failures, and (more pessimistically) motorists learn better how to co-opt the test.

The study presented evaluates the sixth two-year period of the established in Atlanta thirteen counties I/M program. The first evaluation review covered the 1997-

1998 years and the second evaluation covered the years 1999 and 2000. Both of these studies compared the Atlanta inspected fleet with an uninspected fleet in Augusta and Macon. The advent of the Atlanta Clean Fuels program required that the third evaluation, covering the years 2001 and 2002, incorporate a second reference area just outside of the thirteen county I/M area to account for these fuel differences. However, limited measurements in these areas during the third evaluation period resulted in evaluation uncertainties greater than desired and the CAFE program measurement program was modified to dramatically increase the number of vehicles measured in these areas, at the expense of reducing measurements in the I/M area. Table B.2 summarizes results from all six reviews.

As discussed earlier, the changes in the reference areas (and the incorporation of NOx measurements into the analysis) used in these evaluations makes direct comparisons between the first two (1998 and 2000) and the latter four (2002, 2004, 2006 and 2008) evaluations difficult. However, the first group and latter groups may be compared with each other.

The reference method is not without its limitations, however. Selecting a comparable non-program fleet is a challenging task. While differences in fleet age and car/truck composition are relatively easy to account for between I/M and non-I/M fleets, discrepancies in maintenance trends, socioeconomic conditions and vehicle quality are difficult to discern. However, the emissions differences illustrated in Figure B.2 and Figure B.3 make a compelling case that both the I/M and clean fuels program have a significant and positive impact on motor vehicle emissions in the Atlanta area.

Further, Table B.2 illustrates that these emissions differences are durable to the extent that similar results have been observed over an extended period. Table B.2 also reveals a significant difference between NOx benefits achieved for cars and for light duty trucks. While it is tempting to suggest that this is a systematic problem within the program, this may not be the case. The popularity of SUVs during the late 1990's and the early 2000's means that the 2008, 2006 and 2004 (and to a lesser extent the 2002) evaluations of truck emissions are dominated by vehicles for which low emissions reductions are expected. Likewise, for the same reason, the car estimates in 2008 represent, on average, an older fleet than for earlier evaluations. We may therefore hypothesize that differences between "car" and "truck" efficiencies may be an age effect in combination with, or instead of, an intrinsic difference in the effectiveness of the program for these two classes of vehicles. Establishing this relationship will be the subject for future studies.

Table B.2 I/M Effectiveness Estimated for 1998, 2000, 2002, 2004, 2006 and 2008 Measurement Years.

Estimated IM Effectiveness		**Cars**	**Light Trucks**
1998 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	87%	75%
	NOx	NA	NA
2000 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	84%	84%
	NOx	NA	NA
2002 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	166%	229%
	NOx Low estimation	78%	68%
	NOx High estimation	170%	150%

Estimated IM Effectiveness		Cars	Light Trucks
2004 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	256%	223%
	NOx Low estimation	142%	72%
	NOx High estimation	176%	90%
2006 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	146%	126%
	NOx Low estimation	165%	105%
	NOx High estimation	206%	131%
2008 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	134%	79%
	NOx Low estimation	101%	111%
	NOx High estimation	121%	133%

APPENDIX C. BIENNIAL EVALUATION OF THE EMISSIONS REDUCTION EFFECTIVENESS OF THE ATLANTA VEHICLE INSPECTION AND MAINTENANCE PROGRAM FOR 2005-2006

C.1 Introduction

C.1.1 Overview:

The 1990 Clean Air Act Amendments (CAAA) made sweeping changes in the scope and stringency of vehicle inspection/maintenance (I/M) programs. Driven by persistent growth in vehicle travel and chronic air pollution in the nation's largest metropolitan areas, the legislation requires "enhanced" I/M programs employ advanced testing technologies and procedures as a way to better ensure the operability of vehicle emission control system.[55] The law also requires biennial evaluation of enhanced I/M programs and on road measurement of inspected fleet emissions, but does not link together the two requirements (CAA Title I §182c3C; CAA Title I §182c3Bi; Title I§182c3Ci).

In the absence of an explicit legislative linkage, the National Research Council has recommended that I/M programs be evaluated using on road emissions data collected by remote sensing devices (RSD) (National Research Council, 2001). RSD uses infrared and ultraviolet technology to measure the emissions of in-use vehicles.[56] The NRC report cited several advantages of RSD data for I/M evaluation. First, RSD is a cost-effective source of evaluation data compared with the higher per-vehicle costs of advanced dynamometer

[55] Enhanced I/M programs were required in areas of the United States in serious, severe, or extreme nonattainment of federal ozone standards. Moderate nonattainment areas were required to implement the less rigorous basic I/M programs. Marginal nonattainment areas had no I/M requirement.

[56] Infrared technology is used to measure carbon monoxide and volatile organic compounds. Ultraviolet technology is used to measure nitrogen oxides

testing on a small sample of vehicles, the original evaluation approach recommended by federal regulators. RSD data can also capture trends that cannot be discerned through internal inspection records alone, such as motorists avoiding the program and pre-inspection maintenance behavior. RSD data can also be used for a variety of purposes in addition to I/M evaluation, including mobile source emission inventories, clean-screen programs that exempt low-emission vehicles from subsequent I/M testing, and high-emitter programs that target polluting vehicles for off-cycle inspection and repair.

In response to this growing interest, the U.S. Environmental Protection Agency (EPA) released draft guidance in July 2001 for the use of remote sensing data for I/M program evaluation (U.S. EPA, 2001). The document outlines equipment specifications and measurement procedures along with study design techniques and quality control measures. The document also discusses three methodologies for analyzing remote sensing data to determine I/M program effectiveness. *The comprehensive method* compares the onroad emissions of the vehicle fleet before and after scheduled I/M testing. The step *method* compares inspected with uninspected model year emissions during the first year of a new or upgraded I/M program. *The reference method* compares the emissions of the vehicle fleet located in an I/M area with that of a distantly located non-I/M area.

This paper employs the reference method to evaluate the enhanced I/M program of Atlanta, GA. This major metropolitan area in the southeastern United States is home to

thirteen counties in “serious” nonattainment of the federal ozone standard.[57] The Atlanta enhanced I/M program was implemented in October 1996 in this thirteen-county area, replacing a basic I/M program that had been operating in four of the thirteen counties since the early 1980s. We estimate the effectiveness of the new I/M program by comparing the RSD emissions of a sample of its inspected vehicles with that of a sample of vehicles registered in the Georgia cities of Augusta and Macon. The latter areas have demographics, climate and fleet characteristics similar to Atlanta, but do not operate an I/M program. The emissions difference in the inspected Atlanta and uninspected Augusta/Macon vehicle fleets are then compared with that predicted by the commonly used EPA MOBILE6.2 computer model. Viewing model-predicted emissions differences in inspected and uninspected vehicles as the Atlanta I/M program goal and observed onroad emission differences as actual program performance, we estimate I/M effectiveness as the ratio of these two numbers.

This section provides background on I/M programs, including an overview of I/M program operations, and a history of I/M programs in Atlanta. The second section reviews current enhanced I/M evaluation approaches (including the RSD methods outlined in recent EPA guidance) and their respective strengths and weaknesses. The third section describes

[57] The federal ozone standard is 0.12 ppm averaged on an hourly basis and 0.08 ppm averaged over an eight-hour basis.. Ozone concentration is one of the six National Ambient Air Quality Standards set by EPA to protect public health. (The remainder include carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, particulate matter-10, particular matter-2.5, and lead.) There are five levels of nonattainment for these pollutants, ranging from marginal to extreme, which are determined by the number of times air monitoring stations in an area detect pollutant levels above federal standards during a specific timeperiod.

reference method data sources and methodology. The fourth section reports the results of the reference method. The fifth section discusses the results, with particular attention paid to their caveats. The sixth section presents our conclusions about the reference method results for Atlanta and how it compares with previous I/M evaluations using remote sensing data and reference method.

C.1.2 Vehicle Inspection/Maintenance Programs:

Vehicle inspection/maintenance (I/M) programs seek, first and foremost, to ensure the effectiveness of vehicle emission control systems. The inspection process, which applies to light-duty vehicles of a certain age, involves scheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls.[58] Inspections can be provided by decentralized test-and-repair networks, which allow service stations and automotive repair shops to perform emissions tests and repair failed vehicles, or by centralized test-only networks, in which a limited number of centrally operated facilities perform testing as the sole service. Depending on program design, the test may be performed annually or biennially, while the vehicle idles or is placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of on road driving conditions.

[58] The model year vehicles subject to testing vary across I/M programs.

Motorists must repair failed vehicles, comprising the maintenance component of the program. Vehicles with repair costs above a set amount may qualify for a waiver -- an exemption from further repair and testing -- provided that attempted repairs show some emissions improvement and are not triggered by tampering. Compliance is typically verified through the presence of a vehicle windshield sticker received after passing the test or through the vehicle registration process that requires an emissions certificate.

C.1.3 Inspection/Maintenance in Atlanta:

Atlanta's first I/M program was established in 1981, covering the three ozone nonattainment area counties of Fulton, Cobb and DeKalb. The fast-growing Gwinnett County was added in 1986. The program was implemented through a decentralized test-and-repair network which allowed repair shops, service stations and automobile dealers to perform emission inspections and emissions-related repairs. Testing was originally required for the latest ten model year vehicles, but was expanded in 1986 to include the latest twelve model years. To receive an emissions compliance certificate, cars were required to pass an idle emissions test and an inspection of the catalyst, air pump and fuel inlet restrictor for evidence of tampering. Owners of failing cars that spent more than $50 for repairs qualified for a waiver and an emissions certificate, so long as repairs were not due to tampering and showed some emissions improvement. Owners of cars that failed the tampering inspection were required to obtain repairs to bring their emissions into compliance regardless of cost.

In response to the 1990 Clean Air Act Amendments (CAAA), the Georgia legislature revisited emissions testing in 1992[59]. This legislation enabled the Georgia Department of Natural Resources (GDNR) to upgrade Georgia's I/M program to an "enhanced" program, bringing it into compliance with the 1990 CAAA and new federal I/M federal regulations. This enhanced version of the program received limited implementation in October 1996[60], with emission inspections required only for those vehicles migrating to the Atlanta I/M program area. The new program commenced in January 1997, with biennial emissions testing required of all vehicles from the 1975 model year up until two years of age. The new program also spanned the 13-county nonattainment area, incorporating nine new counties that were not subject to the previous basic I/M program.

After the first two-years of operation several changes had been made to the program: the program began to require vehicles over six years of age to undergo the more rigorous Acceleration Simulation Mode (ASM) testing in October 1998, while in the first two-years period all vehicles were subjects of two-speed idle test (TSI)[61]. The primary difference between ASM and TSI testing is the approximation of real-world driving conditions, i.e., placing the engine under load. While the emissions inspector depresses the accelerator to

[59] 1992 Georgia Air Quality Act, Article 2: Motor Vehicle Emissions Inspection and Maintenance Act (OCGA Section 12-9-40 et seq.).

[60] October 1996 was chosen as the soonest possible start-up date after the previous basic I/M program, which operated during a January-to-April vehicle registration "season." Vehicle registration is now conducted year-round in Georgia, as is enhanced emissions testing.

[61] two-speed idle (TSI) testing procedure that measures emissions under idle and a 2500 RPM engine speed

achieve 25 miles per hour (MPH), ASM testing places the vehicle's drive wheels on a treadmill-like dynamometer that applies a 25 percent load on the vehicle engine. The latter approach is more representative of actual driving conditions than an idle test. Vehicles that failed emissions testing were required to be brought into compliance by repair. Owners of covered vehicles in the 13-county ozone nonattainment area were required to show proof of a passing emissions inspection, a waiver, or proof that they qualify for an exemption in order to register their vehicle.

This report concentrates on the fifth two-year period of inspection and maintenance program operation in Atlanta and covers years 2005 and 2006. By this time certain significant changes have been made to the Atlanta enhanced I/M program. The waiver limit had been increased several times. In 2001, testing frequency changed from biennial to annual; the requirement to inspect vehicles back to 1975 model years was replaced with the requirement to inspect the latest 25 model years; and the exemption from testing of the newest two model years was changed to exemption of the newest three model years.

C.2 Enhanced I/M Programs Evaluation Methods

Three types of data currently dominate the evaluation of enhanced I/M programs: I/M records, which document the results of each inspection; roadside pullovers, which administer emissions tests to vehicles of randomly selected willing motorists; and remote sensing data, which measures on road vehicle emissions. This section reviews evaluations employing each data type, along with the strengths and weakness of each.

C.2.1 Emissions Inspection Records:

The most common source of biennial evaluation data comes from emissions test records generated by I/M programs.[62] I/M test records provide a cost-effective source of evaluation data because they are routinely generated and easily accessible. Because I/M records cover the entire inspected vehicle population, statistical conclusion validity is generally not an issue: evaluators can control for a variety of vehicle characteristics that influence emissions. The availability of odometer data in most I/M records is also advantageous, enabling evaluators to control for the influence of mileage on emissions. A final advantage stems from inspection/maintenance protocols, which are designed to correlate with the Federal Test Procedure[63] and to facilitate quality control.

However, I/M records suffer from weaknesses that limit their reliability as the sole indicator of program performance. Chief among these is the inability to parcel out fraudulent testing behavior, particularly when inspectors substitute clean-emitting vehicles for unrepaired high-emitting ones on the retest (Wenzel et al., 2000). I/M records may also underestimate program effectiveness by missing pre-inspection maintenance performed by some motorists to lower I/M test failure risks. While it is difficult to quantify the impact of

---

[62] Personal conversation with James Lindner of the U.S. Environmental Protection Agency's Office of Transportation and Air Quality, September 26, 2001. Also see http://www.epa.gov/otaq/epg/progeval.htm.

[63] The Federal Test Procedure is an elaborate testing protocol established in the early 1970s to certify manufacturer compliance with the 1970 Clean Air Act-mandated new vehicle emission standards.

such maintenance, it is expected to yield artificially low baseline emissions and thus underestimate program effectiveness. Generally speaking, these weaknesses speak to the role of I/M records as an internal, not an independent, source of evaluation data.

Evaluations employing I/M records also make tradeoffs between internal validity and representativeness of the data. The inspection process employs highly-controlled conditions to ensure that vehicles are measured under consistent circumstances (e.g., engine stress, vehicle speed, and temperature). While these controls reduce confounding influences on emissions, they represent only a fraction of driving conditions that typify onroad driving. Consequently, the ability to extrapolate I/M test emissions to onroad emissions is limited.

To estimate I/M effectiveness, some evaluations calculate the average emissions difference between the initial and final test scores on failing vehicles and assume that the difference is attributable to the I/M program. Three studies used this approach to evaluate different time periods of the Arizona enhanced I/M program. Two of these studies (Wenzel, 1999 and Glover and Brzezinksi, 1999) estimated a 14 percent reduction in carbon monoxide (CO), a 15 percent reduction in hydrocarbons (HC), and a seven percent reduction in nitrogen oxides (NOx). The third study (Ando, et al, 1999), focusing on repaired vehicles, estimated emission reductions of eight, eight and fourteen percent for CO, HC and NOx.

Sierra Research (1998) also compared initial and final emission results for failed vehicles in AirCare, the Canadian Vancouver/British Columbia emissions testing program.

This study estimated I/M emission reductions of 13 percent CO, 9 percent HC, and 4 percent NOx. Replacing initial emission results of failed inspections with EPA model predictions of an untested fleet's emissions, the researchers estimated 16 percent, 20 percent, and 14 percent emission reductions for CO, HC and NOx. The latter emission differences are thought to be higher than the former because model predictions, as opposed to initial inspection results, are not influenced by pre-inspection maintenance behavior.

The Colorado enhanced I/M program was twice evaluated using inspection records. The first analysis, comparing final test scores for vehicles inspected in 1997 with the new program's first 2,138 initial inspection test scores in 1995, indicated CO emission reductions in the range of 30 to 34 percent (Environ, 1998). The second analysis compared failed vehicles' initial and final inspection results from 1998 that had been converted to Federal Test Procedure scores. The comparison, which normalized repair benefits to the entire inspected fleet, suggested that CO had been reduced by eight percent and HC by six percent, with NOx increasing by one percent (Office of the State Auditor, 1999). While the study results seem contradictory, they cover different timeframes, make divergent assumptions (about deterioration rates, the fate of vehicles with final failures) and employ different measures in estimating I/M effectiveness.

One weakness in attributing before-after emission differences to I/M is the potential for "regression to the mean" emissions behavior, in which a portion of I/M failures will

register lower emissions on the final inspection without repair.[64] This phenomenon is driven by tremendous emissions test-to-test variability, the presence of vehicles with marginally failing emissions, and variance in environmental conditions favorable to test performance. Without verifying repairs, the emissions differences between initial and final test scores may overestimate program effectiveness.

C.2.2 Roadside Emission Inspections:

Used primarily in California, roadside emissions tests are administered with the aid of law enforcement officers who randomly pull vehicles over and ask motorists to voluntarily submit their vehicles to an emissions inspection. Volunteer license plate numbers are then used to query the I/M program database to determine those vehicles with and without an inspection in the past twelve months. Recently inspected and uninspected vehicle emissions are then compared to estimate the emission reductions due to enhanced I/M. Roadside emissions estimates of 1999 enhanced I/M program effectiveness indicate emission reductions of 13 percent for CO, 14 percent for HC, and 6 percent for NO (California Air Resources Board, 2000).

[64] Regression to the mean occurs when two imperfectly correlated measures are compared for a nonrandom sample. The nonrandom sample is typically drawn from high or low scorers on either measure. Regression to the mean occurs when the sample mean moves towards the population mean in the absence of intervention. In the context of I/M evaluation, this means that certain vehicles failing their initial I/M test will score more closely to the mean of the population on the retest, i.e., register passing emissions, without repair. Regression-to-the mean can also occur in vehicles that pass their initial inspection but would fail a subsequent retest.

As with I/M program data, roadside pullovers enable the collection of odometer data for mileage estimates. In contrast with I/M program data, the spontaneity of roadside inspections preclude fraudulent test results that overestimate effectiveness, as well as pre-inspection maintenance behavior that underestimates program effectiveness. However, because roadside emissions tests employ a portable version of official inspection procedures, they sacrifice real-world driving conditions. Furthermore, the approach is costly and generates limited data, requiring as many as four technicians and one law enforcement officer to measure approximately 25 vehicles per day (Wenzel, et al 2000, p. III-8). Self-selection bias is a risk because the test is voluntary and tends to yield a ten percent refusal rate (Wenzel, et al 2000, p. III-8).[65]

C.2.3 Remote Sensing Data from Onroad Vehicles:

A second source of data for evaluating I/M program effectiveness, the one used in this study, is from remote sensing devices (RSD) that measure the emissions of vehicles while they are being driven. The advantage of in-transit measurement is the ability to observe a vehicle's emissions under typical driving conditions, which cannot be as easily captured by traditional controlled emissions testing procedures. Remote sensors can measure a large number of vehicles, an important attribute given the need to control for

[65] The evidence of such bias is mixed. One recent study that used remote sensing to measure the vehicle emissions of refusals and participants alike found no significant difference between the two groups (Wenzel et al, 2000, pg. III-8), while an earlier similar study found that refusal vehicles had 2.5 times the emissions of volunteer vehicles (Stedman, 1994).

tremendous emissions variability due to vehicle type, age, make and model, and emission control technology. A final advantage stems from the unscheduled nature of the measurement, which precludes pre-inspection and fraudulent maintenance behavior that can occur when motorists (as with I/M tests) know when a measurement will occur.

In contrast with the highly controlled parameters of the emissions inspection, the physical circumstances of remote sensing data collection are only approximated through sampling site characteristics (e.g., moderate grades to ensure vehicles operate under only a slight engine load and sampling sites that avoid residential areas to minimize inflated emissions from cold engines). Another drawback is that remote sensors capture a split-second emissions reading that may not reflect a vehicle's typical emissions, making larger samples sizes a requirement to average out random emission fluctuations and to profile emissions aggregated within vehicle type (cars vs. trucks) and model year.

Remote sensing data has been used in three ways to evaluate I/M programs. The first method averages the emissions of vehicles measured before initial and after final I/M testing, with the difference attributed to I/M program effectiveness. Dubbed the "comprehensive method" in recent EPA evaluation guidance, emissions differences can also be generated for various subfleets, such as vehicles initially failing and ultimately passing I/M testing versus failing vehicles that never receiving a final pass. This approach enables a variety of I/M-related analyses, such as deterioration rates of I/M repairs, the influence of pre-I/M repairs on emissions baselines, and a comparison with estimates based on I/M records alone. The major disadvantage to this approach is the enormous volume of onroad data required to measure a representative sample of vehicles before and after I/M

testing. Sample size requirements hinge on the probability of measuring vehicles onroad within a specific timeperiod of I/M testing, a probability that fluctuates with testing frequency and the distribution of sampling throughout the year.

The comprehensive method was used to estimate the effectiveness of the California South Coast Air Basin's enhanced I/M program in 1999 (Wenzel et al., 2000). "Smog Check" I/M records were used to delineate tested from untested vehicles by the existence of an enhanced inspection within the past twelve months.[66] A comparison of these vehicle groups indicates a ten percent reduction in CO, a four percent reduction in HC, and a five percent increase in NOx. An earlier remote sensing study in California in 1996 compared the onroad emissions of 3.5 million vehicles 30 to 90 days before with up to 90 days after their basic I/M test (Klausmeier and Weyn, 1997). For those vehicles that failed their initial smog check and then passed, both CO and HC emission differences registered at 20 percent. Normalizing this result to the entire fleet yielded an estimated nine percent emissions reduction in HC and CO. A third evaluation, of the Arizona enhanced I/M program in 1997, analyzed four million remote sensing measurements on 1.2 million vehicles in the Phoenix I/M area (Wenzel, 1999). The results indicated a seven percent reduction in CO and an 11 percent reduction in HC.

[66] Untested vehicles may have been inspected under the previous basic I/M program more than twelve months ago or they may have had an enhanced inspection after the remote sensing reading.

One weakness in the comprehensive method is the potential seasonal effects that results from the year-round testing required to obtain adequately sized samples. Users of this method have also tended to rely on a few high-volume sites, yielding a large number of repeat vehicles that lower the fraction of unique vehicles that could be reached at a greater number of sites.

A second I/M evaluation approach using remote sensing, known as the Step Method, compares inspected with uninspected vehicles during the first year of a new or upgraded program. The uninspected vehicles comprise an internal control group against which to compare the emission reductions of the inspected vehicles. Because the method applies to the early phases of a new or improved program, it can be used only once to assess program effectiveness.

A remote sensing study of the Colorado Enhanced I/M program compared odd (inspected) and even (uninspected) model year vehicles during the end of the first year of a new biennial enhanced I/M program (Stedman, et al, 1997). At that point, in program history, all odd model year vehicles should have been inspected, whereas all even model year vehicles had no reason to be inspected. This timing rendered even model year vehicles the untested control group against which to compare the odd model year vehicle emissions. The comparison of odd and even model year emissions suggested that Colorado's enhanced I/M program had reduced CO between five and nine percent, while HC and NO showed no improvement.

Three factors limit the generalizability of the Colorado study results to enhanced I/M program effectiveness. Remote sensing took place in a single location, which avoids any confounding socioeconomic or physical influences at different sites but limits generalizability to the overall fleet. Furthermore, vehicles traveling past the remote sensing site were decelerating, which does not represent typical driving conditions and is not the optimal condition for measuring carbon monoxide (Environ, 1998, p. 2-19). A third limitation was that the study measured vehicles transitioning from an annual basic I/M program to an enhanced I/M program, rendering it an evaluation of incremental program effectiveness and not a complete estimate of enhanced I/M program performance.

A third approach using remote sensing data (the one used in this study) compares the onroad emissions of vehicles registered in an I/M area to that of vehicles registered in non-I/M areas. The non-I/M area serves as a surrogate untested fleet. The validity of this approach relies on the selection of a non-I/M area comparable in fleet age, a well-documented contributor to vehicle emissions; climate, which can accelerate emission control equipment deterioration; and demographics, which influences the age, quality, and maintenance of the vehicle fleet. This approach was originally applied to the basic I/M program operating in four counties of the thirteen-county Atlanta ozone nonattainment area, with the nine nonattainment counties without I/M comprising the untested fleet. The analysis indicates that car and truck emissions for CO were 15 and ten percent higher, respectively, in the uninspected nine-county fleet than in the inspected four-county basic I/M fleet. The study is limited by its inability to control for differences in mileage and socioeconomic conditions between the two vehicle fleets.

C.3 I/M Program Evaluation Components

This study employs an I/M program evaluation method that compares the on road emissions differences *observed* in inspected and uninspected vehicles with the same emissions differences *predicted* by a U.S. Environmental Protection Agency mobile emissions model. The model-predicted emissions difference represents the goal of the I/M program, a reasonable assumption given that states use the model to generate the emission reduction credit received for automobile emissions testing programs. The emissions difference observed in on road inspected and uninspected vehicles is assumed to reflect I/M program performance, an assumption rendered plausible only by the comparability of the inspected and uninspected fleets. We will devote attention in the next section to answering the comparability question.

This section describes the collection of data used in the evaluation. It details the Continuous Atlanta Fleet Evaluation (CAFE), the remote sensing study of on road Georgia vehicles that provides on road emissions data of inspected and uninspected vehicles. The MOBILE6, EPA's recommended emissions model, from which we extracted predicted emission factors, is also discussed. The last section outlines the algorithm that combines data from CAFE and MOBILE6 to generate effectiveness estimates for the Atlanta enhanced I/M program.

C.3.1 On-Road Emissions Data:

The Continuous Atlanta Fleet Evaluation (CAFE) provides the on road emissions data used to represent, *inter alia*, Atlanta enhanced I/M program performance. CAFE uses

remote sensing devices to measure annually the emissions of approximately 380,000 in-use vehicles in the 13-county I/M program area, as well as two cities located more than 75 miles from Atlanta that do not require vehicle emissions testing.[67] The study is an ongoing effort started in 1993 to collect vehicle emissions data for assessing a variety of trends, including fleet turnover, emission control deterioration, and socioeconomic impacts of mobile source control strategies.

RSD measures the emissions of passing vehicles remotely and unobtrusively so motorists are minimally aware of the equipment and do not alter their natural driving behavior. To that end, the remote sensing instrumentation is housed in a van parked on the roadside along with a video camera. An infrared light source and its generator are placed on the opposite side of the road or on the median to create a beam of light that traverses the road. When a passing vehicle breaks the beam, it triggers a measurement of hydrocarbons, carbon monoxide, and nitrogen oxides in the exhaust. Simultaneously, a video camera records the vehicle's license plate, which is automatically scanned into the database of emissions measurements.

After data collection, remote sensing measurements are merged with vehicle registration records using the vehicle license plate. The resulting database allows various characteristics of measured vehicles to be identified, including vehicle identification

[67] Augusta is located 136 miles east of Atlanta, whereas Macon is 76 miles south of Atlanta.

number,[68] make, model year, and vehicle type. License plates are also linked with inspection/maintenance records to identify vehicles with prior emission inspections.

RSD sampling sites are selected to ensure physically consistent but demographically diverse characteristics. Single straight lines of traffic with an average 35 mile-per-hour velocity are sought to facilitate single vehicle measurements and speeds that maximize measurement opportunities. Driver behavior and driving maneuvers are also observed at each site to ensure that remote sensing measurements would not be biased high by acceleration or low by coasting. Finally, notations are made during the site visits regarding any obvious or suspected diurnal patterns that exist which affect the traffic volume. If distinct variations are found to exist in sites ultimately selected, sampling times are scheduled to account for those diurnal patterns. U.S. Census tract data and traffic count reports inform the selection of different income ranges and land uses.

The remote sensing sites relevant to this study reside within the 13-county Atlanta I/M program area, 12 Atlanta counties without an I/M program but subject to the Atlanta clean fuel program, as well as the Georgia cities of Augusta and Macon that have neither program. The latter locales do not require emissions testing and thus provided an

[68] Vehicle identification numbers are 17-digit alphanumeric strings that uniquely identify every vehicle manufactured. When decoded, they provide additional characteristics on vehicles. The VIN-decoded data of particular relevance to this research are vehicle type (car, truck, multi-purpose vehicle, van) and model year.

uninspected vehicle fleet to serve as a control group for our previous I/M evaluations. These cities were chosen after a review of census data and registration records revealed them to have characteristics – median household income, population density, and fleet distribution -- most similar to Atlanta than three other Georgia cities considered. But for the reasons which will be explained later for the present analysis we also used as the reference point the data collected from the vehicles registered in 12 counties that surround Atlanta I/M program area.

According to the state regulation, effective April 1, 1999 the sulfur content of all gasoline supplied in a 25 Atlanta region[69] shall not exceed a seasonal average of 150 ppm (by weight) and, effective April 1, 2001, a per-gallon cap of 500 ppm (by weight)[70]. This rule made vehicle operational conditions in Atlanta 13-county nonattainment area and Augusta-Macon significantly unequal. Since there is no mechanism to separate benefits received from the usage of low sulfur gasoline and emission reductions due to I/M program, the usage of Augusta-Macon fleet as a control group for I/M program evaluation became questionable. In the effort to eliminate the fuel effect, the data collected from the vehicles registered in twelve counties[71] that are not subject to the I/M program but receive the same

[69] 25-county Atlanta region include 13-county I/M program area and 12 additional counties without I/M program: Barrow, Bartow, Butts, Carroll, Dawson, Hall, Haralson, Jackson, Newton, Pickens, Spalding, Walton.

[70] Rules for Air Quality Control Chapter 391-3-1, July 20, 2005.
http://www.gaepd.org/Files_PDF/rules/rules_exist/391-3-1.pdf

[71] Eight of these twelve counties are in the Atlanta Metropolitan Statistical Area and thus are considered “Atlanta-area” counties.

fuel as Atlanta 13-counties I/M program area have been used in the present analysis as a reference point. The data collected on Augusta-Macon sites represents combined benefits from both I/M and GA fuel programs.

C.3.2 Predicted Emission Factors:

We used MOBILE6.2, EPA's recommended computer model for estimation of mobile emission factors, to predict emissions differences in inspected and uninspected vehicles.

C.3.3 Evaluation Algorithm:

We estimated Atlanta enhanced I/M program effectiveness by comparing EPA model-predicted emission differences with observed emission differences in inspected and uninspected vehicles. The comparison yields a percentage that represents the proportion of expected emission reductions actually achieved by the program. The formula for estimating I/M effectiveness is as follows:

$$\text{Effectiveness}=\frac{\sum_{ij}[(O_{n_{ij}}-O_{m_{ij}})/O_{n_{ij}}](P_{n_{ij}})(C_{ij})(VMT_{ij})}{\sum_{ij}(P_{n_{ij}}-P_{m_{ij}})(C_{ij})(VMT_{ij})}$$

where: O_m and O_n are the average onroad emissions observed for a particular model year and vehicle type for I/M program and non-program vehicles, respectively; P_m and P_n are the model-estimated emission factors for I/M program and non-program vehicles for a given model year a vehicle type; C_{ij} is the fraction of the Atlanta fleet of that model year

and vehicle type observed by CAFE; and VMT_{ij} is the average annual vehicle-miles-traveled by model year and vehicle type in the I/M program area.

The formula normalizes predicted and observed emissions differences in I/M program and non-I/M program vehicles by model year to the on road fleet fraction and average annual mileage of that model year. This exercise enables the different units of measurement between on road and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form.

C.4 Analysis

This section reports the results of the reference method for evaluating the Atlanta enhanced I/M program during its third two years of operation. The evaluation uses remote sensing emissions data collected in 2006 and emission factors predicted for the 2006 fleet by an EPA computer model. The 2006 calendar year represents the end of the fifth full cycle of enhanced IM testing and by this time all vehicles should have been inspected under new annual program.

Because the reference method involves direct comparisons between on road data and EPA's MOBILE6.2 model, we restrict the data in several ways to obtain an "apples-to-apples" comparison. First, only 1982 to 2003 model year cars and trucks are included in the analysis. The 2004, 2005 and 2006 model years are not included since these vehicles were exempt from testing in 2006.

The second data restriction is the use of only vehicles registered in the thirteen Atlanta counties of the I/M program area as an *Inspected* fleet while some of the inspected vehicles could move to non I/M areas due to natural migration (such as change of ownership, etc.).

C.4.1 Data Overview:

The remote sensing data used by this evaluation were collected at twenty eight (28) Atlanta I/M program area sites and seven non-program area sites in Augusta and Macon.[72] Measurements in the I/M program area were conducted from January to December 2006, while the non-program area measurements were collected over 18 days in March, June, July, August, September and December.

CAFE collected 161,680 measurements from vehicles registered in thirteen-county area with a 2006 inspection. In the non-program areas, 35,499 measurements were collected from vehicles registered in the counties comprising Augusta and Macon and 12,450 measurements from 12 Atlanta counties that are not subject to the I/M program but are included in the Atlanta Clean Fuels program. The measurements in both of these non-I/M areas are substantially greater than the 16,797 and 8,398 measurements collected in these areas respectively as part of the 2002 evaluation[73]. This shift of measurements into the non-I/M areas was designed to improve the overall statistically validity of the reference

[72] I/M program area measurements are made within thirteen counties that comprise the metropolitan Atlanta ozone nonattainment area: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Fulton, Gwinnett, Henry, Forsyth, Paulding, and Rockdale. Non-I/M program measurements include Bibb, Richmond and Columbia counties.

[73] 2002 was the first year when Atlanta 13 counties were compared to two different control points. In following years number of measurements collected at control sites was increased to improve the quality of analysis.

comparison. In both program and non-program areas, we randomly selected one measurement from unique vehicles with multiple readings.

This evaluation of the Atlanta enhanced I/M program relies on measurements of CO and NOx data. The primary reason for focusing on CO over HC is that the former pollutants have a greater signal-to-noise ratio. CO's lower variance is due to its presence in higher concentrations than HC, making it easier to measure by remote sensing devices and less susceptible to weather and driving conditions. In other words, although HC data was collected and analyzed during the study, these data were not the primary focus of the analysis.

C.4.2 Validity of Fleet Comparisons:

Our ability to infer I/M effectiveness from the emission differences in Atlanta and Augusta/Macon vehicles hinges upon the comparability of the three fleets. The inspected Atlanta and uninspected Augusta-Macon vehicle fleets have similar model year distributions, although the inspected Atlanta fleet is slightly newer than the uninspected Augusta-Macon fleet (Figure C.1).

A second issue for the validity of our analysis is whether the Augusta and Macon fleets are similar enough to be combined into one uninspected fleet. Both this and the previous studies had shown that the average CO emissions by model year and vehicle type do not differ significantly between the two uninspected fleets and thus these results are combined.

The third issue is the ability of the Augusta/Macon fleet to serve as a reference for the estimation of I/M effectiveness in Atlanta. While this was a reasonable assumption for the first two evaluations, the introduction of low sulfur gasoline initiative in 25 Atlanta counties changed this. Vehicles registered in Atlanta 13-county area were thus able to seize benefits from two different emission control programs while vehicles from Augusta and Macon have neither. Thus for purposes of this evaluation, the Augusta-Macon fleet was used to analyze a combined effect of I/M and fuel programs and data collected from 12 Atlanta counties that are not subject to I/M testing was used for estimation of I/M contribution alone.

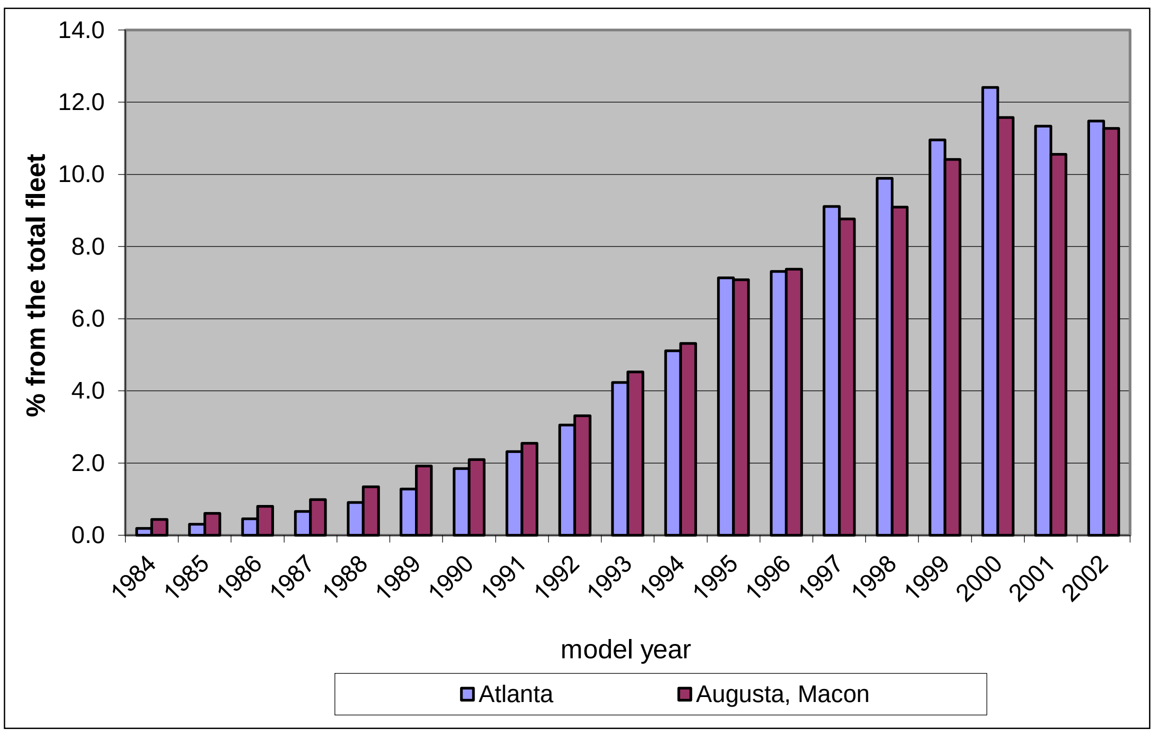

Figure C.1 Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.

C.4.3 Reference Method Results:

The results of the reference method for evaluating the effectiveness of the Atlanta enhanced I/M program are laid out in Table C.1. But first, let us review the methodology for generating the estimates. We calculate the emissions difference in inspected and uninspected cars and trucks by model year and then weight those differences to that model year's annual average mileage and fleet fraction. The exercise is undertaken separately for predicted emissions factors and on road emissions data. The weighted emissions differences in each category are then summed over all model years. The weighted value based on onroad emissions data becomes the numerator, whereas the weight value based on predicted emission factors becomes the denominator. Dividing the numerator by the denominator yields the percentage of expected emissions differences actually achieved in inspected and uninspected vehicles. The results of this exercise indicate that the Atlanta enhanced I/M program captures 146 percent of CO reductions for cars and 126 percent for trucks compared to those predicted by EPA.

Table C.1 Effectiveness of Atlanta I/M Program and Fuel Program.

	Atlanta 13-counties inspected fleet vs. Augusta-Macon uninspected fleet	
	Cars	**Trucks (LDT2)**
CO	146%	126%
NOx, lower estimate	165%	105%
NOx, higher estimate	206%	131%

Delving into the data comprising these results, Figure C.2 and Figure C.3 compare the CO emissions differences in inspected thirteen-county Atlanta and uninspected Augusta-Macon vehicles measured on road by RSD.

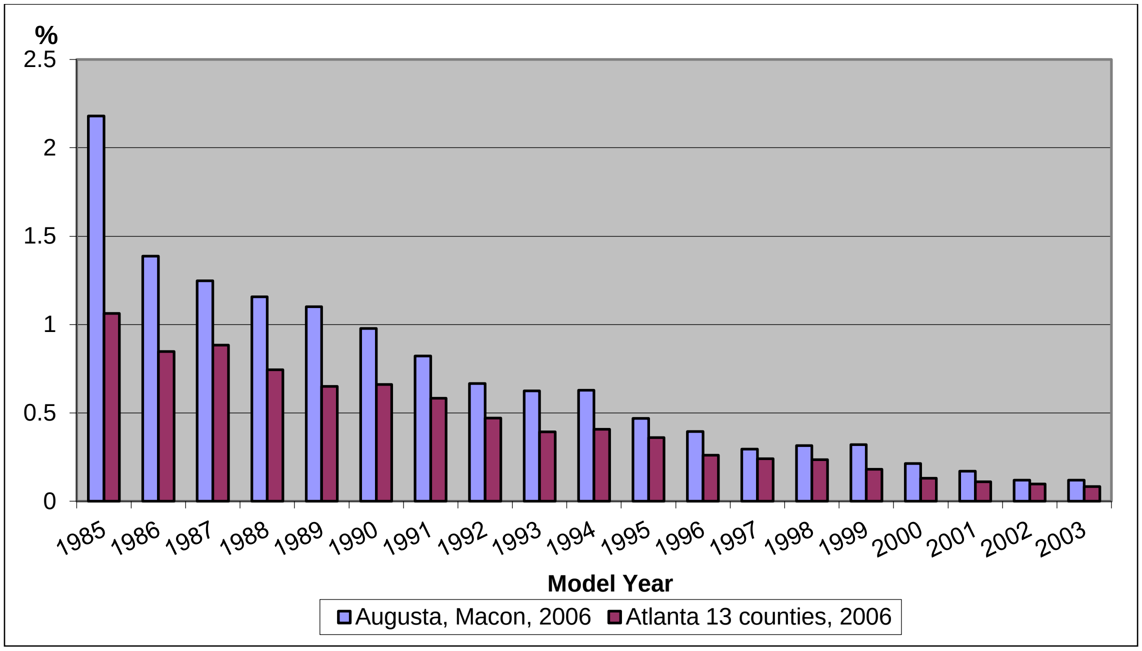

Figure C.2 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.

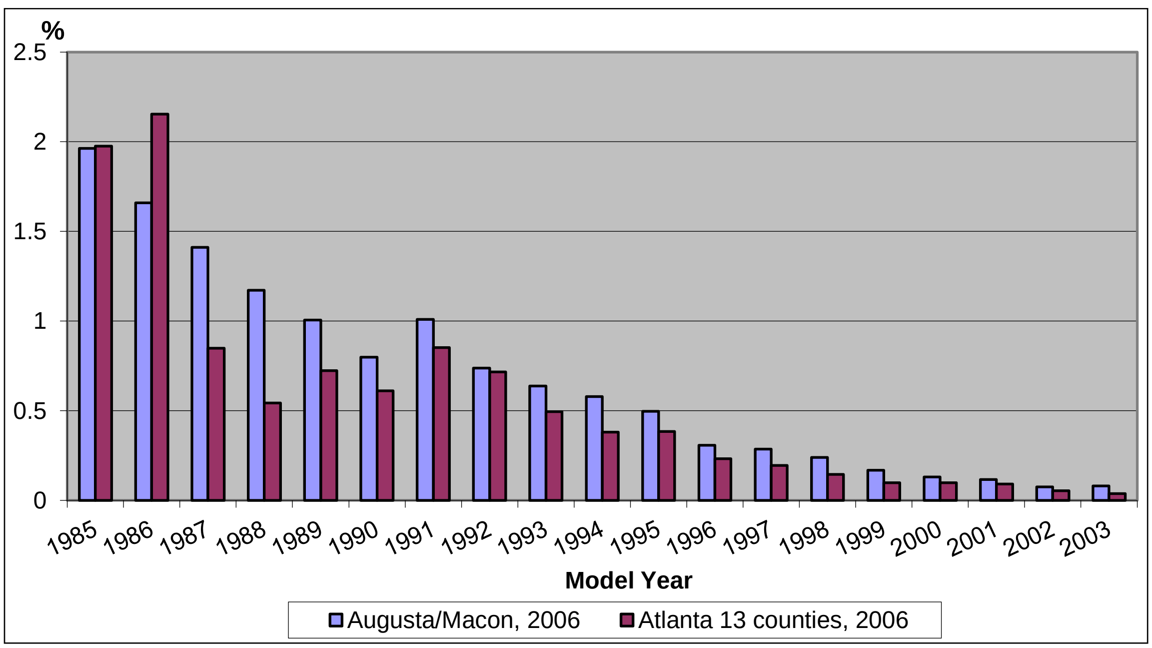

Figure C.3 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. LDT2 Trucks Only.

The on road emission differences for NOx mimic this pattern, although with much larger fluctuations due to additional benefits Atlanta gets from the usage of low sulfur fuel. It is known that the amount of sulfur in the gasoline affects level of NOx exhausted. Figure C.4 illustrates the changes in the average NOx values due to seasonal variations of sulfur level in the gasoline supplied. Therefore additional references are needed to separate I/M air quality benefits and those from low sulfur fuel.

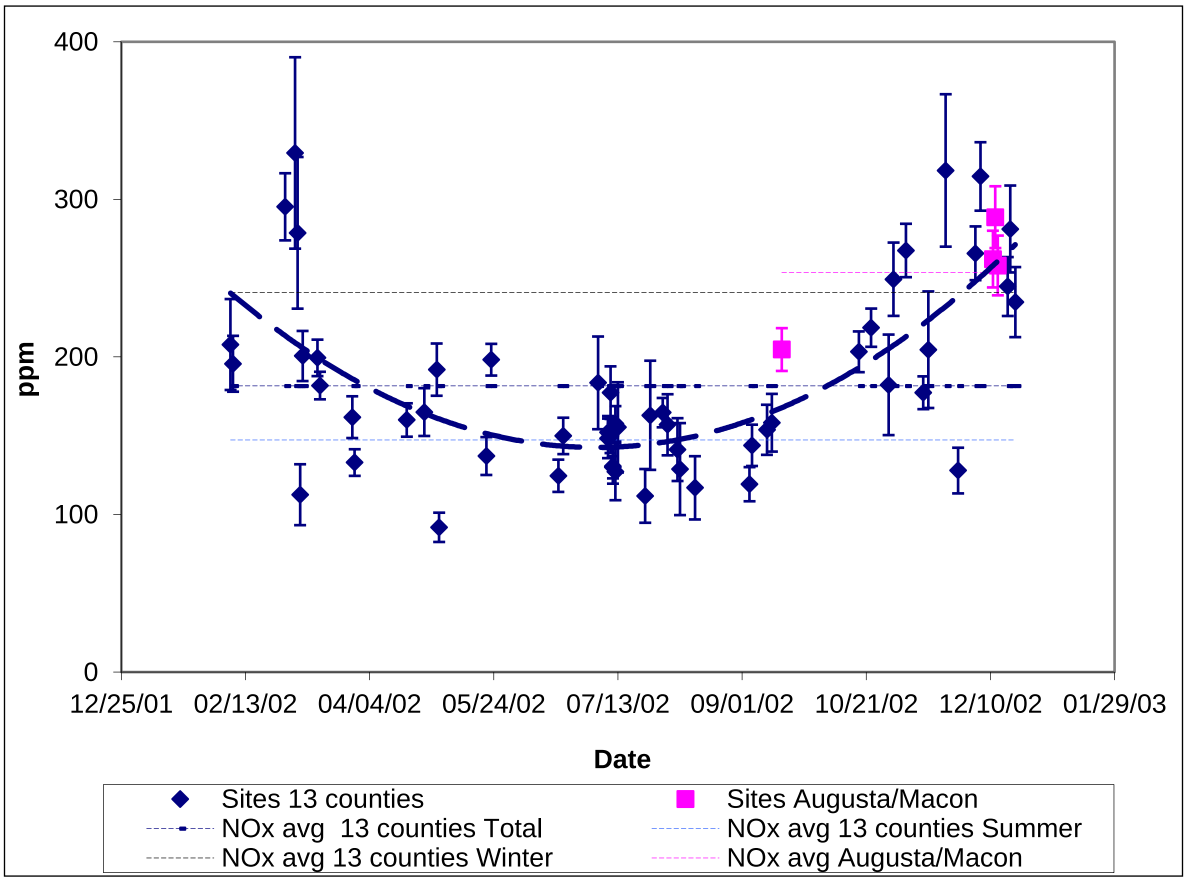

Figure C.4 Seasonal Average Values of NOx. Passenger Cars Only.

The data collected from vehicles recorded in the twelve Atlanta MSA counties that are not subject to I/M but operate on the same fuel was used as one of such references and represents the lower point of our estimated I/M benefits for NOx. The usage of this Atlanta "uninspected" fleet as the solid mark for comparison does not seems justified. First, our evaluation of I/M records has shows that due to local migration between counties about 28% of the vehicles that are not subject to I/M in this area have actually undergone the testing during previous two years. Therefore comparing the Atlanta inspected and "uninspeced" fleets we will tend to underestimate the benefits from I/M program. On the other hand, it is perceived that these same 12 counties may have become a repository for the vehicles that are most likely to fail the I/M test. In other words, people who are trying to avoid I/M testing may be seeking to register their vehicles outside of the I/M area. However, our previous work indicates that number of such vehicles does not exceed 3% from the fleet. Based on the reasoning outlined above we determine the I/M effectiveness for NOx derived from direct comparison of Atlanta 13 counties and Atlanta 12 counties as the lower estimation point assuming that it represents only 75% of actual benefits and higher estimation results from adding an additional 25% benefit for the tested fraction.

C.5 Discussion

Interpreting emissions differences in the Atlanta inspected fleet and Augusta/Macon fleets as combined effect of the enhanced I/M program and fuel programs assumes that we have controlled for all differences in these fleets. This assumption is challenged by the possibility that the Augusta/Macon fleet is composed of higher mileage or poorer quality vehicles than the Atlanta thirteen-county fleet. One source of evidence for mileage differences, the U.S. Department of Transportation data on daily vehicle miles traveled (VMT), suggests that vehicles in Atlanta travel 34 miles per day per capita versus 22 miles for vehicles in Augusta. This information would seem to weaken any hypothesized mileage difference, at least between Augusta and Atlanta. However, because GDOT estimates are based on observed freeway traffic flows that capture out-of-state as well as local vehicles, it is difficult to extrapolate these VMT estimates to the local vehicle fleet. Exclusion of luxury cars from analyzed data sets did not make significant changes in emission patterns therefore the fleet composition differences between Atlanta and Augusta/Macon are negligible.

The comparison of Atlanta thirteen counties inspected fleet with Atlanta uninspected fleet has the same validity issues. Since vehicles that likely to fail testing have the tendency for migration into neighboring counties that are not subject for I/M program we may overestimate its effectiveness. But due to close proximity inspected vehicles also penetrate the noninspected area after change of ownership or under other circumstances which leads to underestimation of I/M benefits.

C.6 Comparing Results with Previous Reviews

The reference method for evaluating vehicle inspection/maintenance programs yields several advantages over other methods using on road remote sensing data. In fact, the reference method could be repeated over time to measure incremental effectiveness as more of the fleet is tested, inspectors become adept at identifying noncompliant vehicles, repair technicians gain experience at repairing emission control failures, and (more pessimistically) motorists learn better how to co-opt the test.

The study presented evaluates the fifth two-year period of the established in Atlanta thirteen counties I/M program. The first evaluation review covered the 1997-1998 years and the second evaluation covered the years 1999 and 2000. Both of these studies compared the Atlanta inspected fleet with an uninspected fleet in Augusta and Macon. The advent of the Atlanta Clean Fuels program required that the third evaluation, covering the years 2001 and 2002, incorporate a second reference area just outside of the thirteen county I/M area to account for these fuel differences. However, limited measurements in these areas during the third evaluation period resulted in evaluation uncertainties greater than desired and the CAFE program measurement program was modified to dramatically increase the number of vehicles measured in these areas, at the expense of reducing measurements in the I/M area. Table C.2 summarizes results from all fife reviews.

As discussed earlier, the changes in the reference areas (and the incorporation of NOx measurements into the analysis) used in these evaluations makes direct comparisons between the first two (1998 and 2000) and the latter three (2002, 2004 and 2006) evaluations difficult. However, the first group and latter groups may be compared with each other.

The reference method is not without its limitations, however. Selecting a comparable non-program fleet is a challenging task. While differences in fleet age and car/truck composition are relatively easy to account for between I/M and non-I/M fleets, discrepancies in maintenance trends, socioeconomic conditions and vehicle quality are difficult to discern. However, the emissions differences illustrated in Figure C.2 and Figure C.3 make a compelling case that both the I/M and clean fuels program have a significant and positive impact on motor vehicle emissions in the Atlanta area.

Further, Table C.2 illustrates that these emissions differences are durable to the extent that similar results have been observed over an extended period. Table C.2 also reveals a significant difference between NOx benefits achieved for cars and for light duty trucks. While it is tempting to suggest that this is a systematic problem within the program, this may not be the case. The popularity of SUVs during the late 1990's and the early 2000's means that the 2006 and 2004 (and to a lesser extent the 2002) evaluations of truck emissions are dominated by vehicles for which low emissions reductions are expected. Likewise, for the same reason, the car estimates in 2006 represent, on average, an older fleet than for earlier evaluations. We may therefore hypothesize that differences between "car" and "truck" efficiencies may be an age effect in combination with, or instead of, an intrinsic difference in the effectiveness of the program for these two classes of vehicles. Establishing this relationship will be the subject for future studies.

Table C.2 I/M Effectiveness Estimated for 1998, 2000, 2002, 2004 and 2006 Measurement Years.

Estimated IM Effectiveness		**Cars**	**Light Trucks**
1998 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	87%	75%
	NOx	NA	NA
2000 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	84%	84%
	NOx	NA	NA
2002 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	166%	229%
	NOx Low estimation	78%	68%
	NOx High estimation	170%	150%
2004 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	256%	223%
	NOx Low estimation	142%	72%
	NOx High estimation	176%	90%
2006 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	146%	126%
	NOx Low estimation	165%	105%
	NOx High estimation	206%	131%

APPENDIX D. BIENNIAL EVALUATION OF VEHICLE INSPECTION/MAINTENANCE PROGRAM USING ON-ROAD EMISSION DATA FOR 2003-2004

D.1 Introduction

The 1990 Clean Air Act Amendments (CAAA) made sweeping changes in the scope and stringency of vehicle inspection/maintenance (I/M) programs. Driven by persistent growth in vehicle travel and chronic air pollution in the nation's largest metropolitan areas, the legislation requires "enhanced" I/M programs employ advanced testing technologies and procedures as a way to better ensure the operability of vehicle emission control system.[74] The law also requires biennial evaluation of enhanced I/M programs and onroad measurement of inspected fleet emissions, but does not link together the two requirements (CAA Title I §182c3C; CAA Title I §182c3Bi; Title I§182c3Ci).

In the absence of an explicit legislative linkage, a National Research Council report has recommended that I/M programs be evaluated using onroad emissions data collected by remote sensing devices (RSD) (National Research Council, 2001). RSD uses infrared and ultraviolet technology to measure the emissions of in-use vehicles.[75] The NRC report cited several advantages of RSD data for I/M evaluation. First, RSD is a cost-effective source of evaluation data compared with the higher per-vehicle costs of advanced dynamometer testing on a small sample of vehicles, the original evaluation approach

---

[74] Enhanced I/M programs are required in areas of the United States in serious, severe, or extreme nonattainment of federal ozone standards. Moderate nonattainment areas are required to implement the less rigorous basic I/M programs. Marginal nonattainment areas have no I/M requirement.

[75] Infrared technology is used to measure carbon monoxide and volatile organic compounds. Ultraviolet technology is used to measure nitrogen oxides.

recommended by federal regulators. RSD data can also capture trends that cannot be discerned through internal inspection records alone, such as motorists avoiding the program and pre-inspection maintenance behavior. RSD data can also be used for a variety of purposes in addition to I/M evaluation, including mobile source emission inventories, clean-screen programs that exempt low-emission vehicles from subsequent I/M testing, and high-emitter programs that target polluting vehicles for off-cycle inspection and repair.

In response to this growing interest, the U.S. Environmental Protection Agency (EPA) released draft guidance in July 2001 for the use of remote sensing data for I/M program evaluation (U.S. EPA, 2001). The document outlines equipment specifications and measurement procedures along with study design techniques and quality control measures. The document also discusses three methodologies for analyzing remote sensing data to determine I/M program effectiveness. *The comprehensive method* compares the onroad emissions of the vehicle fleet before and after scheduled I/M testing. *The step method* compares inspected with uninspected model year emissions during the first year of a new or upgraded I/M program. *The reference method* compares the emissions of the vehicle fleet located in an I/M area with that of a distantly located non-I/M area.

This paper employs the reference method to evaluate the enhanced I/M program of Atlanta, GA. This major metropolitan area in the southeastern United States is home to

thirteen counties in "serious" nonattainment of the federal ozone standard.76 The Atlanta enhanced I/M program was implemented in October 1996 in this thirteen-county area, replacing a basic I/M program that had been operating in four of the thirteen counties since the early 1980s. We estimate the effectiveness of the new I/M program by comparing the RSD emissions of a sample of its inspected vehicles with that of a sample of vehicles registered in the Georgia cities of Augusta and Macon. The latter areas have demographics, climate and fleet characteristics similar to Atlanta, but do not operate an I/M program. The emissions difference in the inspected Atlanta and uninspected Augusta/Macon vehicle fleets are then compared with that predicted by the commonly used EPA MOBILE6.2 computer model. Viewing model-predicted emissions differences in inspected and uninspected vehicles as the Atlanta I/M program goal and observed onroad emission differences as actual program performance, we estimate I/M effectiveness as the ratio of the two numbers.

This section provides background on I/M programs, including an overview of I/M program operations, and a history of I/M programs in Atlanta. The second section reviews current enhanced I/M evaluation approaches (including the RSD methods outlined in recent EPA guidance) and their respective strengths and weaknesses. The third section describes

---

[76] The federal ozone standard is 0.12 ppm averaged on an hourly basis. Ozone concentration is one of the six National Ambient Air Quality Standards set by EPA to protect public health. (The remainder include carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, particulate matter-10, particular matter-2.5, and lead.) There are five levels of nonattainment for these pollutants, ranging from marginal to extreme, which are determined by the number of times air monitoring stations in an area detect pollutant levels above federal standards during a specific timeperiod.

reference method data sources and methodology. The fourth section reports the results of the reference method. The fifth section discusses the results, with particular attention paid to their caveats. The sixth section presents our conclusions about the reference method results for Atlanta and how it compares with previous I/M evaluation using remote sensing data and reference method.

D.1.1 Vehicle Inspection/Maintenance Programs

Vehicle inspection/maintenance (I/M) programs seek, first and foremost, to ensure the effectiveness of vehicle emission control systems. The inspection process, which applies to light-duty vehicles of a certain age, involves scheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls.[77] Inspections can be provided by decentralized test-and-repair networks, which allow service stations and automotive repair shops to perform emissions tests and repair failed vehicles, or by centralized test-only networks, in which a limited number of centrally operated facilities perform testing as the sole service. Depending on program design, the test may be performed annually or biennially, while the vehicle idles or is placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of onroad driving conditions.

---

[77] The model year vehicles subject to testing vary across I/M programs.

Motorists must repair failed vehicles, comprising the maintenance component of the program. Vehicles with repair costs above a set amount may qualify for a waiver -- an exemption from further repair and testing -- provided that attempted repairs show some emissions improvement and are not triggered by tampering. Compliance is typically verified through the presence of a vehicle windshield sticker received after passing the test or through the vehicle registration process that requires an emissions certificate.

D.1.2 Inspection/Maintenance In Atlanta

Atlanta's first I/M program were established in 1981, covering the three ozone nonattainment area counties of Fulton, Cobb and DeKalb. The fast-growing Gwinnett County was added in 1986. The program was implemented through a decentralized test-and-repair network which allowed repair shops, service stations and automobile dealers to perform emission inspections and emissions-related repairs. Testing was originally required for the latest ten model year vehicles, but was expanded in 1986 to include the latest twelve model years. To receive an emissions compliance certificate, cars were required to pass an idle emissions test and an inspection of the catalyst, air pump and fuel inlet restrictor for evidence of tampering. Owners of failing cars that spent more than $50 for repairs qualified for a waiver and an emissions certificate, so long as repairs were not due to tampering and showed some emissions improvement. Owners of cars that failed the tampering inspection were required to obtain repairs to bring their emissions into compliance regardless of cost.

In response to the 1990 Clean Air Act Amendments (CAAA), the Georgia legislature revisited emissions testing in 1992[78]. This legislation enabled the Georgia Department of Natural Resources (GDNR) to upgrade Georgia's I/M program to an "enhanced" program, bringing it into compliance with the 1990 CAAA and new federal I/M federal regulations. This enhanced version of the program received limited implementation in October 1996[79], with emission inspections required only for those vehicles migrating to the Atlanta I/M program area. The new program commenced in January 1997, with biennial emissions testing required of all vehicles from the 1975 model year to two years of age. The new program also spanned the 13-county nonattainment area, incorporating nine new counties that were not subject to the previous basic I/M program.

After the first two-years of operation several changes had been made to the program: the program began to require vehicles over six years of age to undergo the more rigorous Acceleration Simulation Mode (ASM) testing in October 1998, while in the first two-years period all vehicles were subjects of two-speed idle test (TSI)[80]. The primary difference between ASM and TSI testing is the approximation of real-world driving conditions, i.e., placing the engine under load. While the emissions inspector depresses the accelerator to

---

[78] 1992 Georgia Air Quality Act, Article 2: Motor Vehicle Emissions Inspection and Maintenance Act (OCGA Section 12-9-40 et seq.).

[79] October 1996 was chosen as the soonest possible start-up date after the previous basic I/M program, which operated during a January-to-April vehicle registration "season." Vehicle registration is now conducted year-round in Georgia, as is enhanced emissions testing.

[80] two-speed idle (TSI) testing procedure that measures emissions under idle and a 2500 RPM engine speed

achieve 25 miles per hour (MPH), ASM testing places the vehicle's drive wheels on a treadmill-like dynamometer that applies a 25 percent load on the vehicle engine. The latter approach is more representative of actual driving conditions than an idle test. Vehicles that failed emissions testing were required to be brought into compliance by repair. Owners of covered vehicles in the 13-county ozone nonattainment area were required to show proof of a passing emissions inspection, a waiver, or proof that they qualify for an exemption in order to register their vehicle.

This report concentrates on the fourth two-year period of inspection and maintenance program operation in Atlanta and covers years 2003 and 2004. By this time certain significant changes have been made to the Atlanta enhanced I/M program. The waiver limit increased several times. In 2001, testing frequency was changed from biennial to annual; the requirement to inspect vehicles back to 1975 model years was replaced with the requirement to inspect the latest 25 model years; and the exemption from testing of the newest two model years was changed to exemption of the newest three model years.

D.2 Enhanced I/M Programs Evaluations

Three types of data currently dominate the evaluation of enhanced I/M programs: I/M records, which document the results of each inspection; roadside pullovers, which administer emissions tests to vehicles of randomly selected willing motorists; and remote sensing data, which measures onroad vehicle emissions. This section reviews evaluations employing each data type, along with the strengths and weakness of each.

D.2.1. Emissions Inspection Records

The most common source of biennial evaluation data comes from emissions test records generated by I/M programs.[81] I/M test records provide a cost-effective source of evaluation data because they are routinely generated and easily accessible. Because I/M records cover the entire inspected vehicle population, statistical conclusion validity is generally not an issue: evaluators can control for a variety of vehicle characteristics that influence emissions. The availability of odometer data in most I/M records is also advantageous, enabling evaluators to control for the influence of mileage on emissions. A final advantage stems from inspection/maintenance protocols, which are designed to correlate with the Federal Test Procedure[82] and to facilitate quality control.

However, I/M records suffer from weaknesses that limit their reliability as the sole indicator of program performance. Chief among these is the inability to parcel out fraudulent testing behavior, particularly when inspectors substitute clean-emitting vehicles for unrepaired high-emitting ones on the retest (Wenzel et al., 2000). I/M records may also underestimate program effectiveness by missing pre-inspection maintenance performed by some motorists to lower I/M test failure risks. While it is difficult to quantify the impact of

---

[81] Personal conversation with James Lindner of the U.S. Environmental Protection Agency's Office of Transportation and Air Quality, September 26, 2001. Also see http://www.epa.gov/otaq/epg/progeval.htm.

[82] The Federal Test Procedure is an elaborate testing protocol established in the early 1970s to certify manufacturer compliance with the 1970 Clean Air Act-mandated new vehicle emission standards.

such maintenance, it is expected to yield artificially low baseline emissions and thus underestimate program effectiveness. Generally speaking, these weaknesses speak to the role of I/M records as an internal, not an independent, source of evaluation data.

Evaluations employing I/M records also make tradeoffs between internal validity and representativeness of the data. The inspection process employs highly-controlled conditions to ensure that vehicles are measured under consistent circumstances (e.g., engine stress, vehicle speed, and temperature). While these controls reduce confounding influences on emissions, they represent only a fraction of driving conditions that typify onroad driving. Consequently, the ability to extrapolate I/M test emissions to onroad emissions is limited.

To estimate I/M effectiveness, some evaluations calculate the average emissions difference between the initial and final test scores on failing vehicles and assume that the difference is attributable to the I/M program. Three studies used this approach to evaluate different time periods of the Arizona enhanced I/M program. Two of these studies (Wenzel, 1999 and Glover and Brzezinksi, 1999) estimated a 14 percent reduction in carbon monoxide (CO), a 15 percent reduction in hydrocarbons (HC), and a seven percent reduction in nitrogen oxides (NOx). The third study (Ando, et al, 1999), focusing on repaired vehicles, estimated emission reductions of eight, eight and fourteen percent for CO, HC and NOx.

Sierra Research (1998) also compared initial and final emission results for failed vehicles in AirCare, the Canadian Vancouver/British Columbia emissions testing program.

This study estimated I/M emission reductions of 13 percent CO, 9 percent HC, and 4 percent NOx. Replacing initial emission results of failed inspections with EPA model predictions of an untested fleet's emissions the researchers estimated 16 percent, 20 percent, and 14 percent emission reductions for CO, HC and NOx. The latter emission differences are thought to be higher than the former because model predictions, as opposed to initial inspection results, are not influenced by pre-inspection maintenance behavior.

The Colorado enhanced I/M program was twice evaluated using inspection records. The first analysis, comparing final test scores for vehicles inspected in 1997 with the new program's first 2,138 initial inspection test scores in 1995, indicated CO emission reductions in the range of 30 to 34 percent (Environ, 1998). The second analysis compared failed vehicles' initial and final inspection results from 1998 that had been converted to Federal Test Procedure scores. The comparison, which normalized repair benefits to the entire inspected fleet, suggested that CO had been reduced by eight percent and HC by six percent, with NOx increasing by one percent (Office of the State Auditor, 1999). While the study results seem contradictory, they cover different timeframes, make divergent assumptions (about deterioration rates, the fate of vehicles with final failures) and employ different measures in estimating I/M effectiveness.

One weakness in attributing before-after emission differences to I/M is the potential for "regression to the mean" emissions behavior, in which a portion of I/M failures will

register lower emissions on the final inspection without repair.[83] This phenomenon is driven by tremendous emissions test-to-test variability, the presence of vehicles with marginally failing emissions, and variance in environmental conditions favorable to test performance. Without verifying repairs, the emissions differences between initial and final test scores may overestimate program effectiveness.

D.2.2 Roadside Emission Inspections

Used primarily in California, roadside emissions tests are administered with the aid of law enforcement officers who randomly pull vehicles over and ask motorists to voluntarily submit their vehicles to an emissions inspection. Volunteer license plate numbers are then used to query the I/M program database to determine those vehicles with and without an inspection in the past twelve months. Recently inspected and uninspected vehicle emissions are then compared to estimate the emission reductions due to enhanced I/M. Roadside emissions estimates of 1999 enhanced I/M program effectiveness indicate emission reductions of 13 percent for CO, 14 percent for HC, and 6 percent for NO (California Air Resources Board, 2000).

[83] Regression to the mean occurs when two imperfectly correlated measures are compared for a nonrandom sample. The nonrandom sample is typically drawn from high or low scorers on either measure. Regression to the mean occurs when the sample mean moves towards the population mean in the absence of intervention. In the context of I/M evaluation, this means that certain vehicles failing their initial I/M test will score more closely to the mean of the population on the retest, i.e., register passing emissions, without repair. Regression-to-the mean can also occur in vehicles that pass their initial inspection but would fail a subsequent retest.

As with I/M program data, roadside pullovers enable the collection of odometer data for mileage estimates. In contrast with I/M program data, the spontaneity of roadside inspections preclude fraudulent test results that overestimate effectiveness, as well as pre-inspection maintenance behavior that underestimates program effectiveness. However, because roadside emissions tests employ a portable version of official inspection procedures, they sacrifice real-world driving conditions. Furthermore, the approach is costly and generates limited data, requiring as many as four technicians and one law enforcement officer to measure approximately 25 vehicles per day (Wenzel, et al 2000, p. III-8). Self-selection bias is a risk because the test is voluntary and tends to yield a ten percent refusal rate (Wenzel, et al 2000, p. III-8).[84]

D.2.3 Remote Sensing Data from Onroad Vehicles

A second source of data for evaluating I/M program effectiveness, the one used in this study, is from remote sensing devices (RSD) that measure the emissions of vehicles while they are being driven. The advantage of in-transit measurement is the ability to observe a vehicle's emissions under typical driving conditions, which cannot be as easily captured by traditional controlled emissions testing procedures. Remote sensors can measure a large number of vehicles, an important attribute given the need to control for

[84] The evidence of such bias is mixed. One recent study that used remote sensing to measure the vehicle emissions of refusals and participants alike found no significant difference between the two groups (Wenzel et al, 2000, pg. III-8), while an earlier similar study found that refusal vehicles had 2.5 times the emissions of volunteer vehicles (Stedman, 1994).

tremendous emissions variability due to vehicle type, age, make and model, and emission control technology. A final advantage stems from the unscheduled nature of the measurement, which precludes pre-inspection and fraudulent maintenance behavior that can occur when motorists (as with I/M tests) know when a measurement will occur.

In contrast with the highly controlled parameters of the emissions inspection, the physical circumstances of remote sensing data collection are only approximated through sampling site characteristics (e.g., moderate grades to ensure vehicles operate under only a slight engine load and sampling sites that avoid residential areas to minimize inflated emissions from cold engines). Another drawback is that remote sensors capture a split-second emissions reading that may not reflect a vehicle’s typical emissions, making larger samples sizes a requirement to average out random emission fluctuations and to profile emissions aggregated within vehicle type (cars vs. trucks) and model year.

Remote sensing data has been used in three ways to evaluate I/M programs. The first method averages the emissions of vehicles measured before initial and after final I/M testing, with the difference attributed to I/M program effectiveness. Dubbed the “comprehensive method” in recent EPA evaluation guidance, emissions differences can also be generated for various subfleets, such as vehicles initially failing and ultimately passing I/M testing versus failing vehicles that never receiving a final pass. This approach enables a variety of I/M-related analyses, such as deterioration rates of I/M repairs, the influence of pre-I/M repairs on emissions baselines, and a comparison with estimates based on I/M records alone. The major disadvantage to this approach is the enormous volume of onroad data required to measure a representative sample of vehicles before and after I/M

testing. Sample size requirements hinge on the probability of measuring vehicles onroad within a specific timeperiod of I/M testing, a probability that fluctuates with testing frequency and the distribution of sampling throughout the year.

The comprehensive method was used to estimate the effectiveness of the California South Coast Air Basin's enhanced I/M program in 1999 (Wenzel et al., 2000). "Smog Check" I/M records were used to delineate tested from untested vehicles by the existence of an enhanced inspection within the past twelve months.[85] A comparison of these vehicle groups indicates a ten percent reduction in CO, a four percent reduction in HC, and a five percent increase in NOx. An earlier remote sensing study in California in 1996 compared the onroad emissions of 3.5 million vehicles 30 to 90 days before with up to 90 days after their basic I/M test (Klausmeier and Weyn, 1997). For those vehicles that failed their initial smog check and then passed, both CO and HC emission differences registered at 20 percent. Normalizing this result to the entire fleet yielded an estimated nine percent emissions reduction in HC and CO. A third evaluation, of the Arizona enhanced I/M program in 1997, analyzed four million remote sensing measurements on 1.2 million vehicles in the Phoenix I/M area (Wenzel, 1999). The results indicated a seven percent reduction in CO and an 11 percent reduction in HC.

[85] Untested vehicles may have been inspected under the previous basic I/M program more than twelve months ago or they may have had an enhanced inspection after the remote sensing reading.

One weakness in the comprehensive method is the potential seasonal effects that results from the year-round testing required to obtain adequately sized samples. Users of this method have also tended to rely on a few high-volume sites, yielding a large number of repeat vehicles that lower the fraction of unique vehicles that could be reached at a greater number of sites.

A second I/M evaluation approach using remote sensing, known as the Step Method, compares inspected with uninspected vehicles during the first year of a new or upgraded program. The uninspected vehicles comprise an internal control group against which to compare the emission reductions of the inspected vehicles. Because the method applies to the early phases of a new or improved program, it can be used only once to assess program effectiveness.

A remote sensing study of the Colorado Enhanced I/M program compared odd (inspected) and even (uninspected) model year vehicles during the end of the first year of a new biennial enhanced I/M program (Stedman, et al, 1997). At that point, in program history, all odd model year vehicles should have been inspected, whereas all even model year vehicles had no reason to be inspected. This timing rendered even model year vehicles the untested control group against which to compare the odd model year vehicle emissions. The comparison of odd and even model year emissions suggested that Colorado's enhanced I/M program had reduced CO between five and nine percent, while HC and NO showed no improvement.

Three factors limit the generalizability of the Colorado study results to enhanced I/M program effectiveness. Remote sensing took place in a single location, which avoids any confounding socioeconomic or physical influences at different sites but limits generalizability to the overall fleet. Furthermore, vehicles traveling past the remote sensing site were decelerating, which does not represent typical driving conditions and is not the optimal condition for measuring carbon monoxide (Environ, 1998, p. 2-19). A third limitation was that the study measured vehicles transitioning from an annual basic I/M program to an enhanced I/M program, rendering it an evaluation of incremental program effectiveness and not a complete estimate of enhanced I/M program performance.

A third approach using remote sensing data (the one used in this study) compares the onroad emissions of vehicles registered in an I/M area to that of vehicles registered in non-I/M areas. The non-I/M area serves as a surrogate untested fleet. The validity of this approach relies on the selection of a non-I/M area comparable in fleet age, a well-documented contributor to vehicle emissions; climate, which can accelerate emission control equipment deterioration; and demographics, which influences the age, quality, and maintenance of the vehicle fleet. This approach was originally applied to the basic I/M program operating in four counties of the thirteen-county Atlanta ozone nonattainment area, with the nine nonattainment counties without I/M comprising the untested fleet. The analysis indicates that car and truck emissions for CO were 15 and ten percent higher, respectively, in the uninspected nine-county fleet than in the inspected four-county basic I/M fleet. The study is limited by its inability to control for differences in mileage and socioeconomic conditions between the two vehicle fleets.

D.3 I/M Program Evaluation Components

This study employs an I/M program evaluation method that compares the onroad emissions differences *observed* in inspected and uninspected vehicles with the same emissions differences *predicted* by a U.S. Environmental Protection Agency mobile emissions model. The model-predicted emissions difference represents the goal of the I/M program, a reasonable assumption given that states use the model to generate the emission reduction credit received for automobile emissions testing programs. The emissions difference observed in onroad inspected and uninspected vehicles is assumed to reflect I/M program performance, an assumption rendered plausible only by the comparability of the inspected and uninspected fleets. We will devote attention in the next section to answering the comparability question.

This section describes the collection of data used in the evaluation. It details the Continuous Atlanta Fleet Evaluation (CAFE), the remote sensing study of onroad Georgia vehicles that provides onroad emissions data of inspected and uninspected vehicles. The MOBILE6, EPA's recommended emissions model, from which we extracted predicted emission factors, is also discussed. The last section outlines the algorithm that combines data from CAFE and MOBILE6 to generate effectiveness estimates for the Atlanta enhanced I/M program.

D.3.1 Onroad Emissions Data

The Continuous Atlanta Fleet Evaluation (CAFE) provides the onroad emissions data used to represent, *inter alia*, Atlanta enhanced I/M program performance. CAFE uses

remote sensing devices to measure annually the emissions of approximately 400,000 in-use vehicles in the 13-county I/M program area, as well as two cities located more than 75 miles from Atlanta that do not require vehicle emissions testing.[86] The study is an ongoing effort started in 1993 to collect vehicle emissions data for assessing a variety of trends, including fleet turnover, emission control deterioration, and socioeconomic impacts of mobile source control strategies.

RSD measures the emissions of passing vehicles remotely and unobtrusively so motorists are minimally aware of the equipment and do not alter their natural driving behavior. To that end, the remote sensing instrumentation is housed in a van parked on the roadside along with a videocamera. An infrared light source and its generator are placed on the opposite side of the road or on the median to create a beam of light that traverses the road. When a passing vehicle breaks the beam, it triggers a measurement of hydrocarbons, carbon monoxide, and nitrogen oxides in the exhaust. Simultaneously, a videocamera records the vehicle's license plate, which is automatically scanned into the database of emissions measurements.

After data collection, remote sensing measurements are merged with vehicle registration records using the vehicle license plate. The resulting database allows various characteristics of measured vehicles to be identified, including vehicle identification

[86] Augusta is located 136 miles east of Atlanta, whereas Macon is 76 miles south of Atlanta.

number,[87] make, model year, and vehicle type. License plates are also linked with inspection/maintenance records to identify vehicles with prior emission inspections.

RSD sampling sites are selected to ensure physically consistent but demographically diverse characteristics. Single straight lines of traffic with an average 35 mile-per-hour velocity are sought to facilitate single vehicle measurements and speeds that maximize measurement opportunities. Driver behavior and driving maneuvers are also observed at each site to ensure that remote sensing measurements would not be biased high by acceleration or low by coasting. Finally, notations are made during the site visits regarding any obvious or suspected diurnal patterns that exist which affect the traffic volume. If distinct variations are found to exist in sites ultimately selected, sampling times are scheduled to account for those diurnal patterns. U.S. Census tract data and traffic count reports inform the selection of different income ranges and land uses.

The remote sensing sites relevant to this study reside within the 13-county Atlanta I/M program area, 8 Atlanta counties without I/M program as well as the Georgia cities of Augusta and Macon. The latter locales do not require emissions testing and thus provided an uninspected vehicle fleet to serve as a control group for our previous I/M evaluations. These cities were chosen after a review of census data and registration records revealed

[87] Vehicle identification numbers are 17-digit alphanumeric strings that uniquely identify every vehicle manufactured. When decoded, they provide additional characteristics on vehicles. The VIN-decoded data of particular relevance to this research are vehicle type (car, truck, multi-purpose vehicle, van) and model year.

them to have characteristics – median household income, population density, and fleet distribution -- most similar to Atlanta than three other Georgia cities considered. But for the reasons which will be explained latter for the present analysis we also used as the reference point the data collected from the vehicles registered in 12 counties that surround Atlanta I/M program area.

According to the state regulation, effective April 1, 1999 the sulfur content of all gasoline supplied in a 25 Atlanta region[88] shall not exceed a seasonal average of 150 ppm (by weight) and, effective April 1, 2001, a per-gallon cap of 500 ppm (by weight)[89]. This rule made vehicle operational conditions in Atlanta 13-county nonattainment area and Augusta-Macon significantly unequal. Since there is no mechanism to separate benefits received from the usage of low sulfur gasoline and emission reductions due to I/M program, the usage of Augusta-Macon fleet as a control group for I/M program evaluation became questionable. In the effort to eliminate the fuel effect, the data collected from the vehicles registered in twelve counties that are not subject to the I/M program but receive the same fuel as Atlanta 13-counties I/M program area have been used in the present analysis as an additional reference point. Results derived from comparing Atlanta 13-counties

---

[88] 25-county Atlanta region include 13-county I/M program area and 12 additional counties without I/M program: Barrow, Bartow, Butts, Carroll, Dawson, Hall, Haralson, Jackson, Newton, Pickens, Spalding, Walton.

[89] Rules for Air Quality Control Chapter 391-3-1, July 20, 2005. http://www.gaepd.org/Files_PDF/rules/rules_exist/391-3-1.pdf

measurements with those collected on Augusta-Macon sites represent combined benefits from both I/M and GA fuel programs.

D.3.2 Predicted Emission Factors

We used MOBILE6.2, an EPA's recommended computer model for estimation of mobile emission factors, to predict emissions differences in inspected and uninspected vehicles.

D.3.3 Evaluation Algorithm

We estimated Atlanta enhanced I/M program effectiveness by comparing EPA model-predicted emission differences with observed emission differences in inspected and uninspected vehicles. The comparison yields a percentage that represents the proportion of expected emission reductions actually achieved by the program. The formula for estimating I/M effectiveness is as follows:

$$\text{Effectiveness} = \frac{\sum_{ij}[(O_{n_{ij}} - O_{m_{ij}})/O_{n_{ij}}](P_{n_{ij}})(C_{ij})(VMT_{ij})}{\sum_{ij}(P_{n_{ij}} - P_{m_{ij}})(C_{ij})(VMT_{ij})}$$

where: O_m and O_n are the average onroad emissions observed for a particular model year and vehicle type for I/M program and non-program vehicles, respectively; P_m and P_n are the model-estimated emission factors for I/M program and non-program vehicles for a given model year a vehicle type; C_{ij} is the fraction of the Atlanta fleet of that model year

and vehicle type observed by CAFE; and VMTij is the average annual vehicle-miles traveled by model year and vehicle type in the I/M program area.

The formula normalizes predicted and observed emissions differences in I/M program and non-I/M program vehicles by model year to the onroad fleet fraction and average annual mileage of that model year. This exercise enables the different units of measurement between onroad and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form.

D.4 Analysis

This section reports the results of the reference method for evaluating the Atlanta enhanced I/M program during its third two years of operation. The evaluation uses remote sensing emissions data collected in 2004 and emission factors predicted for the 2004 fleet by an EPA computer model. The 2004 calendar year represents the end of the fourth full cycle of enhanced IM testing and by this time all vehicles should have been inspected under annual program.

Because the reference method involves direct comparisons between onroad data and EPA's MOBILE model, we restrict the data in several ways to obtain an "apples-to-apples" comparison. First, only 1980 to 2002 model year vehicles are included in the analysis. While the Atlanta enhanced I/M program inspected back to the 1978 model year during 2004 low sample sizes among onroad vehicles precluded us from valid statistical analysis on cars manufactured earlier than in 1980. The 2003 and 2004 model years are not included since these vehicles were exempt from testing in 2004.

The second data restriction is the use of only vehicles registered in the thirteen Atlanta counties of the I/M program area as an *Inspected* fleet while some of the inspected vehicles could move to non I/M areas due to natural migration (such as change of ownership, etc.).

D.4.1 Data Overview

The remote sensing data used by this evaluation were collected at thirty-two (32) Atlanta I/M program area sites and five non-program area sites in Augusta and Macon.[90] Measurements in the I/M program area were conducted from January to December 2004, while non-program area readings were collected over four days in February, June, August, and September.

CAFE collected 128,870 measurements from vehicles registered in thirteen-county area with a 2004 inspection. In the non-program areas, 62,771 measurements were collected from vehicles registered in the counties comprising Augusta and Macon and 27,508 measurements from 12 Atlanta counties that are not subject to the I/M program. In both program and non-program areas, we randomly selected one measurement from unique vehicles with multiple readings.

[90] I/M program area measurements are made within thirteen counties that comprise the metropolitan Atlanta ozone nonattainment area: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Fulton, Gwinnett, Henry, Forsyth, Paulding, and Rockdale. Non-I/M program measurements include Bibb, Richmond and Columbia counties.

This evaluation of the Atlanta enhanced I/M program relies on measurements of CO and NOx data. The primary reason for focusing on CO over HC is that the former pollutants have a smaller signal-to-noise ratio. CO's lower variance is due to its presence in higher concentrations in the atmosphere than HC, making it easier to measure by remote sensing devices and less susceptible to weather and driving conditions. In other words, although HC data was collected and analyzed during the study, we did not concentrate on it due to large reading errors.

D.4.2 Validity of Fleet Comparisons

Our ability to infer I/M effectiveness from the emission differences in Atlanta and Augusta/Macon vehicles hinges upon the comparability of the three fleets. The inspected Atlanta and uninspected Augusta-Macon vehicle fleets have similar model year distributions, although the inspected Atlanta fleet is slightly newer than the uninspected Augusta-Macon fleet (Figure D.1).

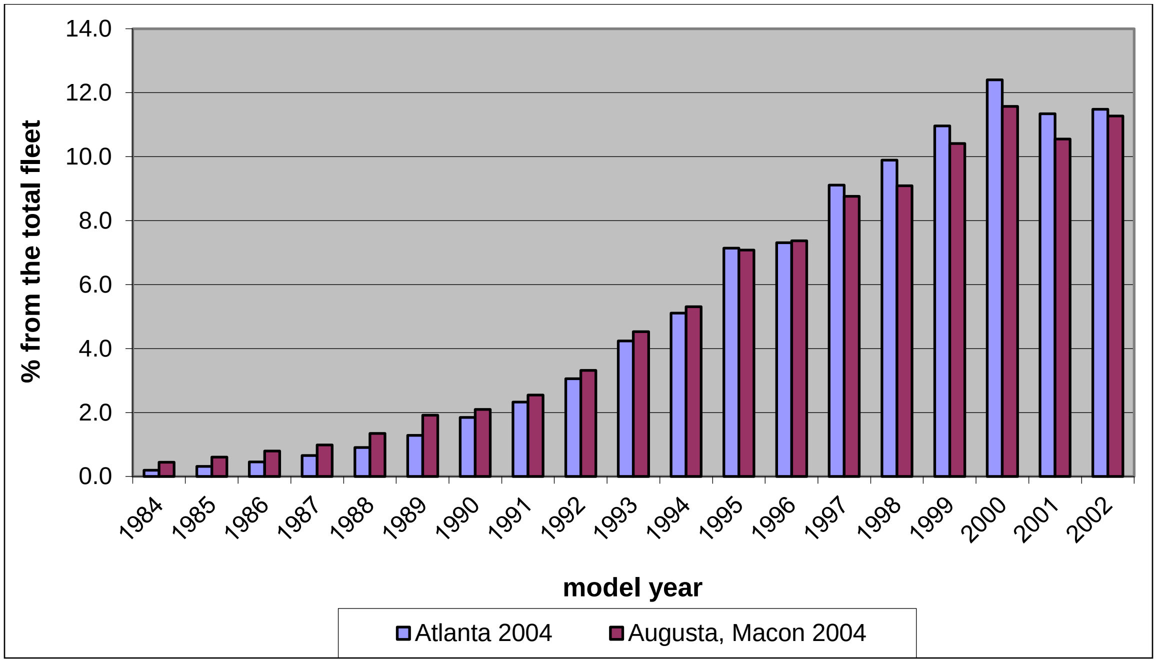

Figure D.1 Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.

A second issue for the validity of our analysis is whether the Augusta and Macon fleets are similar enough to be combined into one uninspected fleet. The previous studies had shown that the average CO emissions by model year and vehicle type do not differ significantly between the two fleets therefore the data collected in both cities can be pooled into one dataset representing uninspected vehicles. As was stated above, introduction of low sulfur gasoline initiative in 25 Atlanta counties made the usage of Augusta and Macon fleets as the only reference point questionable. Vehicles registered in Atlanta 13-county area seize benefits from two different emission control programs while vehicles from Augusta and Macon have neither. Taking into consideration former facts Augusta-Macon fleet was used to analyze a combined effect of I/M and fuel programs and data collected from 12 Atlanta counties that are not subject to I/M testing was used for estimation of I/M contribution.

D.4.3 Reference Method Results

The results of the reference method for evaluating the effectiveness of the Atlanta enhanced I/M program are laid out in Table D.1 (for the thorough examination see Appendix to the present document). But first, let us review the methodology for generating the estimates. We calculate the emissions difference in inspected and uninspected cars and trucks by model year and then weight those differences to that model year's annual average mileage and fleet fraction. The exercise is undertaken separately for predicted emissions factors and onroad emissions data. The weighted emissions differences in each category are then summed over all model years. The weighted value based on onroad emissions data becomes the numerator, whereas the weight value based on predicted emission factors becomes the denominator. Dividing the numerator by the denominator yields the percentage of expected emissions differences actually achieved in inspected and uninspected vehicles. The results of this exercise indicate that the effect of Atlanta enhanced I/M program captures 256 percent of CO reduction for cars and 223 percent for trucks as from predicted by EPA.

Table D.1 Effectiveness of Atlanta I/M Program and Fuel Program.

	Atlanta 13-counties inspected fleet vs. Augusta-Macon uninspected fleet	
	Cars	**Trucks**
CO	256%	223%
NOx, lower estimate	141%	72%
NOx, higher estimate	176%	90%

Delving into the data comprising these results, Figure D.2 and Figure D.3 compare the CO emissions differences in inspected thirteen-county Atlanta and uninspected Augusta-Macon vehicles measured onroad by RSD.

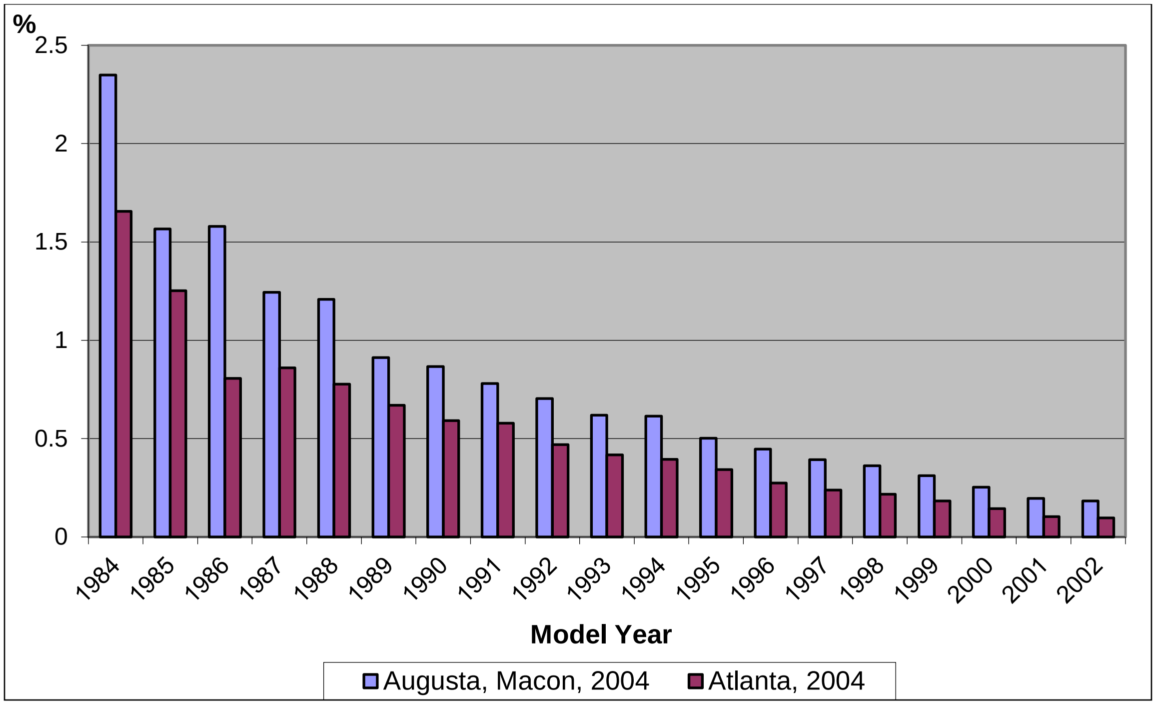

Figure D.2 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.

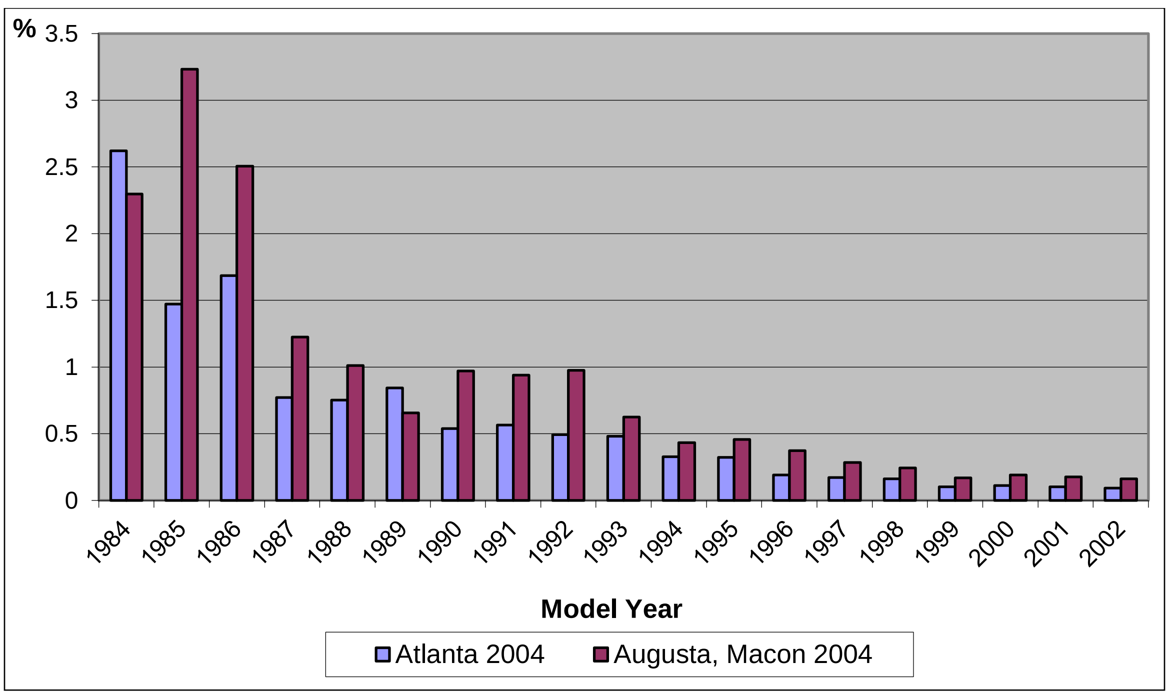

Figure D.3 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Trucks Only.

The onroad emission differences for NOx mimic this pattern, although with much larger fluctuations due to additional benefits Atlanta gets from the usage of low sulfur fuel (see Appendix). It is known that the amount of sulfur in the gasoline affects level of NOx exhausted. Figure D.4 illustrates the changes in the average NOx values due to seasonal variations of sulfur level in the gasoline supplied. Therefore additional references are needed to separate I/M air quality benefits and those from low sulfur fuel.

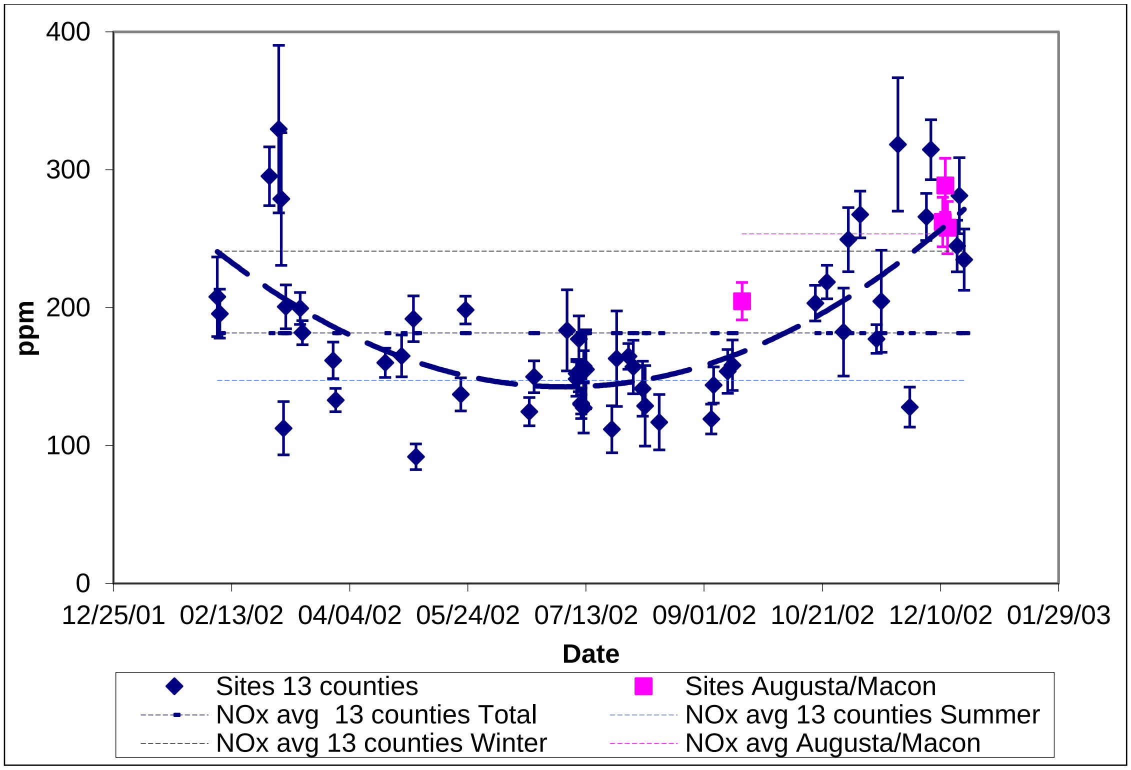

Figure D.4 Seasonal Average Values of NOx. Passenger Cars Only.

The data collected from vehicles registered in twelve Atlanta MSA counties that are not subject to I/M but operate on the same fuel was used as one of such references and represents the lower point of our estimated I/M benefits for NOx. The usage of Atlanta uninspected fleet as the solid mark for comparison does not seems justified. First, our estimation show that due to local migration between counties about 28% of vehicles that are not subject for I/M actually undergone the test during past two years. Therefore by comparing Atlanta inspected and uninspeced fleets we will underestimate benefits from I/M program. On the other hand, it is perceived that Atlanta 12 counties area became a port for the vehicles that are most likely to fail the test. In other words, people who are trying to avoid I/M testing seeking the way to register their vehicles outside of I/M area.

However, our previous work indicates that number of such vehicles does not exceed 3% from the fleet. Based on the reasoning outlined above we settle I/M effectiveness for NOx derived from direct comparison of Atlanta 13 counties and Atlanta 12 counties as the lower estimation point assuming that it represents only 75% of actual benefits and higher estimation results from adding additional 25% (for the thorough analysis see Appendix to present report).

D.5 Discussion

Interpreting emissions differences in the Atlanta inspected fleet and Augusta/Macon fleets as combined effect of the enhanced I/M program and fuel programs assumes that we have controlled for all differences in these fleets. This assumption is challenged by the possibility that the Augusta/Macon fleet is composed of higher mileage or poorer quality vehicles than the Atlanta thirteen-county fleet. One source of evidence for mileage differences, the U.S. Department of Transportation data on daily vehicle miles traveled (VMT), suggests that vehicles in Atlanta travel 34 miles per day per capita versus 22 miles for vehicles in Augusta. This information would seem to weaken any hypothesized mileage difference, at least between Augusta and Atlanta. However, because GDOT estimates are based on observed freeway traffic flows that capture out-of-state as well as local vehicles, it is difficult to extrapolate these VMT estimates to the local vehicle fleet. Exclusion of luxury cars from analyzed data sets did not make significant changes in emission patterns therefore the fleet composition differences between Atlanta and Augusta/Macon are negligible.

The comparison of Atlanta thirteen counties inspected fleet with Atlanta uninspected fleet has the same validity issues. Since vehicles that likely to fail testing have the tendency for migration in neighboring counties that are not subject for I/M program we may overestimate its effectiveness. But due to close proximity inspected vehicles also penetrate the noninspected area after change of ownership or under other circumstances which leads to underestimation of I/M benefits.

D.6 Comparing Results with Previous Reviews.

The reference method for evaluating vehicle inspection/maintenance programs yields several advantages over other methods using onroad remote sensing data. In fact, the reference method could be repeated over time to measure incremental effectiveness as more of the fleet is tested, inspectors become adept at identifying noncompliant vehicles, repair technicians gain experience at repairing emission control failures, and (more pessimistically) motorists learn better how to co-opt the test.

The study presented evaluates the fourth two-year period of the established in Atlanta thirteen counties IM program and focuses on 2003/2004 years. The first evaluation review covered the 1997/1998 years, the second evaluation covered 1999/2000, and the third - 2001/2002. Table D. summarizes results yielded by all four reviews.

Table D.2 IM Effectiveness Estimated for 1998, 2000, 2002, and 2004 Measurement Year.

Estimated IM Effectiveness		Cars	Light Trucks
1998 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	87%	75%
	NOx	NA	NA
2000 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	84%	84%
	NOx	NA	NA
2002 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	166%	229%
	NOx Low estimation	78%	68%
	NOx High estimation	170%	150%
2004 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet and Atlanta uninspected fleet	**CO**	256%	223%
	NOx Low estimation	142%	72%
	NOx High estimation	176%	90%

Unfortunately for the reasons discussed in previous sections of this report results seized by first, second and last two analyses can not be directly compared. In last two studies I/M effectiveness represented by figures derived not only from comparison of inspected Atlanta and uninspected Augusta/Macon fleets but also by using uninspected Atlanta 12 counties fleet as an additional reference point. The comparison of vehicles registered in Atlanta I/M program area and those from Augusta/Macon represents the combined effect of I/M and state fuel programs while evaluation of inspected and uninspected Atlanta fleets correspond only to benefits from I/M program.

As for the Atlanta enhanced I/M program overall, the reference method suggests that the program is reducing onroad emissions, but may not be meeting EPA model predictions for NOx emissions especially for trucks. Future research efforts will include replicating the reference method at different points in place and time to estimate the impact of emission reduction programs changes discussed in this study on onroad fleet emissions.

The reference method is not without its limitations, however. Selecting a comparable non-program fleet is a challenging task, to say the least. While differences in fleet age and car/truck composition are relatively easy to account for between I/M and non-I/M fleets, discrepancies in maintenance trends, socioeconomic conditions and vehicle quality are difficult to discern. Until additional studies of non-I/M fleets shed light on the role of these influences on fleet emissions, equating fleet differences with I/M effectiveness will be a tentative proposition.

APPENDIX E. BIENNIAL EVALUATION OF VEHICLE INSPECTION/MAINTENANCE PROGRAM USING ON-ROAD EMISSION DATA FOR 2001-2002

E.1 Introduction

The 1990 Clean Air Act Amendments (CAAA) made sweeping changes in the scope and stringency of vehicle inspection/maintenance (I/M) programs. Driven by persistent growth in vehicle travel and chronic air pollution in the nation's largest metropolitan areas, the legislation requires "enhanced" I/M programs employ advanced testing technologies and procedures as a way to better ensure the operability of vehicle emission control system.[91] The law also requires biennial evaluation of enhanced I/M programs and onroad measurement of inspected fleet emissions, but does not link together the two requirements (CAA Title I §182c3C; CAA Title I §182c3Bi; Title I§182c3Ci).

In the absence of an explicit legislative linkage, a recent National Research Council report has recommended that I/M programs be evaluated using onroad emissions data collected by remote sensing devices (RSD) (National Research Council, 2001). RSD uses infrared and ultraviolet technology to measure the emissions of in-use vehicles.[92] The NRC report cited several advantages of RSD data for I/M evaluation. First, RSD is a cost-effective source of evaluation data compared with the higher per-vehicle costs of advanced dynamometer testing on a small sample of vehicles, the original evaluation approach

[91] Enhanced I/M programs are required in areas of the United States in serious, severe, or extreme nonattainment of federal ozone standards. Moderate nonattainment areas are required to implement the less rigorous basic I/M programs. Marginal nonattainment areas have no I/M requirement.

[92] Infrared technology is used to measure carbon monoxide and volatile organic compounds. Ultraviolet technology is used to measure nitrogen oxides.

recommended by federal regulators. RSD data can also capture trends that cannot be discerned through internal inspection records alone, such as motorists avoiding the program and pre-inspection maintenance behavior. RSD data can also be used for a variety of purposes in addition to I/M evaluation, including mobile source emission inventories, clean-screen programs that exempt low-emission vehicles from subsequent I/M testing, and high-emitter programs that target polluting vehicles for off-cycle inspection and repair.

In response to this growing interest, the U.S. Environmental Protection Agency (EPA) released draft guidance in July 2001 for the use of remote sensing data for I/M program evaluation (U.S. EPA, 2001). The document outlines equipment specifications and measurement procedures along with study design techniques and quality control measures. The document also discusses three methodologies for analyzing remote sensing data to determine I/M program effectiveness. *The comprehensive method* compares the onroad emissions of the vehicle fleet before and after scheduled I/M testing. *The step method* compares inspected with uninspected model year emissions during the first year of a new or upgraded I/M program. *The reference method* compares the emissions of the vehicle fleet located in an I/M area with that of a distantly located non-I/M area.

This paper employs the reference method to evaluate the enhanced I/M program of Atlanta, GA. This major metropolitan area in the southeastern United States is home to

thirteen counties in “serious” nonattainment of the federal ozone standard.[93] The Atlanta enhanced I/M program was implemented in October 1996 in this thirteen-county area, replacing a basic I/M program that had been operating in four of the thirteen counties since the early 1980s. We estimate the effectiveness of the new I/M program by comparing the RSD emissions of a sample of its inspected vehicles with that of a sample of vehicles registered in the Georgia cities of Augusta and Macon. The latter areas have demographics, climate and fleet characteristics similar to Atlanta, but do not operate an I/M program. The emissions difference in the inspected Atlanta and uninspected Augusta/Macon vehicle fleets are then compared with that predicted by the commonly used EPA MOBILE6.2 computer model. Viewing model-predicted emissions differences in inspected and uninspected vehicles as the Atlanta I/M program goal and observed onroad emission differences as actual program performance, we estimate I/M effectiveness as the ratio of the two numbers.

This section provides background on I/M programs, including an overview of I/M program operations, and a history of I/M programs in Atlanta. The second section reviews current enhanced I/M evaluation approaches (including the RSD methods outlined in recent EPA guidance) and their respective strengths and weaknesses. The third section describes

[93] The federal ozone standard is 0.12 ppm averaged on an hourly basis. Ozone concentration is one of the six National Ambient Air Quality Standards set by EPA to protect public health. (The remainder include carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, particulate matter-10, particular matter-2.5, and lead.) There are five levels of nonattainment for these pollutants, ranging from marginal to extreme, which are determined by the number of times air monitoring stations in an area detect pollutant levels above federal standards during a specific timeperiod.

reference method data sources and methodology. The fourth section reports the results of the reference method. The fifth section discusses the results, with particular attention paid to their caveats. The sixth section presents our conclusions about the reference method results for Atlanta and how it compares with previous I/M evaluation using remote sensing data and reference method.

E.1.1 Vehicle Inspection/Maintenance Programs

Vehicle inspection/maintenance (I/M) programs seek, first and foremost, to ensure the effectiveness of vehicle emission control systems. The inspection process, which applies to light-duty vehicles of a certain age, involves scheduled testing of a vehicle's tailpipe and evaporative emissions to determine the effectiveness of its emission controls.[94] Inspections can be provided by decentralized test-and-repair networks, which allow service stations and automotive repair shops to perform emissions tests and repair failed vehicles, or by centralized test-only networks, in which a limited number of centrally operated facilities perform testing as the sole service. Depending on program design, the test may be performed annually or biennially, while the vehicle idles or is placed on a treadmill-like dynamometer that induces slight acceleration to mimic the engine stress of onroad driving conditions.

[94] The model year vehicles subject to testing vary across I/M programs.

Motorists must repair failed vehicles, comprising the maintenance component of the program. Vehicles with repair costs above a set amount may qualify for a waiver -- an exemption from further repair and testing -- provided that attempted repairs show some emissions improvement and are not triggered by tampering. Compliance is typically verified through the presence of a vehicle windshield sticker received after passing the test or through the vehicle registration process that requires an emissions certificate.

E.1.2 Inspection/Maintenance In Atlanta

Atlanta's first I/M program was established in 1981, covering the three ozone nonattainment area counties of Fulton, Cobb and DeKalb. The fast-growing Gwinnett County was added in 1986. The program was implemented through a decentralized test-and-repair network which allowed repair shops, service stations and automobile dealers to perform emission inspections and emissions-related repairs. Testing was originally required for the latest ten model year vehicles, but was expanded in 1986 to include the latest twelve model years. To receive an emissions compliance certificate, cars were required to pass an idle emissions test and an inspection of the catalyst, air pump and fuel inlet restrictor for evidence of tampering. Owners of failing cars that spent more than $50 for repairs qualified for a waiver and an emissions certificate, so long as repairs were not due to tampering and showed some emissions improvement. Owners of cars that failed the tampering inspection were required to obtain repairs to bring their emissions into compliance regardless of cost.

In response to the 1990 Clean Air Act Amendments (CAAA), the Georgia legislature revisited emissions testing in 1992[95]. This legislation enabled the Georgia Department of Natural Resources (GDNR) to upgrade Georgia's I/M program to an “enhanced” program, bringing it into compliance with the 1990 CAAA and new federal I/M federal regulations. This enhanced version of the program received limited implementation in October 1996[96], with emission inspections required only for those vehicles migrating to the Atlanta I/M program area. The new program commenced in January 1997, with biennial emissions testing required of all vehicles from the 1975 model year to two years of age. The new program also spanned the 13-county nonattainment area, incorporating nine new counties that were not subject to the previous basic I/M program.

After the first two-years of operation several changes had been made to the program: the program began to require vehicles over six years of age to undergo the more rigorous Acceleration Simulation Mode (ASM) testing in October 1998, while in the first two-years period all vehicles were subjects of two-speed idle test (TSI)[97]. The primary difference between ASM and TSI testing is the approximation of real-world driving conditions, i.e., placing the engine under load. While the emissions inspector depresses the accelerator to

---

[95] 1992 Georgia Air Quality Act, Article 2: Motor Vehicle Emissions Inspection and Maintenance Act (OCGA Section 12-9-40 et seq.).

[96] October 1996 was chosen as the soonest possible start-up date after the previous basic I/M program, which operated during a January-to-April vehicle registration “season.” Vehicle registration is now conducted year-round in Georgia, as is enhanced emissions testing.

[97] two-speed idle (TSI) testing procedure that measures emissions under idle and a 2500 RPM engine speed

achieve 25 miles per hour (MPH), ASM testing places the vehicle's drive wheels on a treadmill-like dynamometer that applies a 25 percent load on the vehicle engine. The latter approach is more representative of actual driving conditions than an idle test. Vehicles that failed emissions testing were required to be brought into compliance by repair. Owners of covered vehicles in the 13-county ozone nonattainment area were required to show proof of a passing emissions inspection, a waiver, or proof that they qualify for an exemption in order to register their vehicle.

This report concentrates on the third two-year period of inspection and maintenance program operation in Atlanta and covers years 2001 and 2002. By this time certain significant changes have been made to the Atlanta enhanced I/M program. The waiver limit increased in January 2000 to $608, which represents $450 plus the consumer price index and fulfilled EPA requirements. In 2001, testing frequency changed from biennial to annual; the requirement to inspect vehicles back to 1975 model years was replaced with the requirement to inspect the latest 25 model years; and the exemption from testing of the newest two model years was changed to exemption of the newest three model years.

E.2 Enhanced I/M Programs Evaluations

Three types of data currently dominate the evaluation of enhanced I/M programs: I/M records, which document the results of each inspection; roadside pullovers, which administer emissions tests to vehicles of randomly selected willing motorists; and remote sensing data, which measures onroad vehicle emissions. This section reviews evaluations employing each data type, along with the strengths and weakness of each.

E.2.1 Emissions Inspection Records

The most common source of biennial evaluation data comes from emissions test records generated by I/M programs.[98] I/M test records provide a cost-effective source of evaluation data because they are routinely generated and easily accessible. Because I/M records cover the entire inspected vehicle population, statistical conclusion validity is generally not an issue: evaluators can control for a variety of vehicle characteristics that influence emissions. The availability of odometer data in most I/M records is also advantageous, enabling evaluators to control for the influence of mileage on emissions. A final advantage stems from inspection/maintenance protocols, which are designed to correlate with the Federal Test Procedure[99] and to facilitate quality control.

However, I/M records suffer from weaknesses that limit their reliability as the sole indicator of program performance. Chief among these is the inability to parcel out fraudulent testing behavior, particularly when inspectors substitute clean-emitting vehicles for unrepaired high-emitting ones on the retest (Wenzel et al., 2000). I/M records may also underestimate program effectiveness by missing pre-inspection maintenance performed by some motorists to lower I/M test failure risks. While it is difficult to quantify the impact of

---

[98] Personal conversation with James Lindner of the U.S. Environmental Protection Agency's Office of Transportation and Air Quality, September 26, 2001. Also see http://www.epa.gov/otaq/epg/progeval.htm.

[99] The Federal Test Procedure is an elaborate testing protocol established in the early 1970s to certify manufacturer compliance with the 1970 Clean Air Act-mandated new vehicle emission standards.

such maintenance, it is expected to yield artificially low baseline emissions and thus underestimate program effectiveness. Generally speaking, these weaknesses speak to the role of I/M records as an internal, not an independent, source of evaluation data.

Evaluations employing I/M records also make tradeoffs between internal validity and representativeness of the data. The inspection process employs highly-controlled conditions to ensure that vehicles are measured under consistent circumstances (e.g., engine stress, vehicle speed, and temperature). While these controls reduce confounding influences on emissions, they represent only a fraction of driving conditions that typify onroad driving. Consequently, the ability to extrapolate I/M test emissions to onroad emissions is limited.

To estimate I/M effectiveness, some evaluations calculate the average emissions difference between the initial and final test scores on failing vehicles and assume that the difference is attributable to the I/M program. Three studies used this approach to evaluate different time periods of the Arizona enhanced I/M program. Two of these studies (Wenzel, 1999 and Glover and Brzezinksi, 1999) estimated a 14 percent reduction in carbon monoxide (CO), a 15 percent reduction in hydrocarbons (HC), and a seven percent reduction in nitrogen oxides (NOx). The third study (Ando, et al, 1999), focusing on repaired vehicles, estimated emission reductions of eight, eight and fourteen percent for CO, HC and NOx.

Sierra Research (1998) also compared initial and final emission results for failed vehicles in AirCare, the Canadian Vancouver/British Columbia emissions testing program.

This study estimated I/M emission reductions of 13 percent CO, 9 percent HC, and 4 percent NOx. Replacing initial emission results of failed inspections with EPA model predictions of an untested fleet's emissions, the researchers estimated 16 percent, 20 percent, and 14 percent emission reductions for CO, HC and NOx. The latter emission differences are thought to be higher than the former because model predictions, as opposed to initial inspection results, are not influenced by pre-inspection maintenance behavior.

The Colorado enhanced I/M program was twice evaluated using inspection records. The first analysis, comparing final test scores for vehicles inspected in 1997 with the new program's first 2,138 initial inspection test scores in 1995, indicated CO emission reductions in the range of 30 to 34 percent (Environ, 1998). The second analysis compared failed vehicles' initial and final inspection results from 1998 that had been converted to Federal Test Procedure scores. The comparison, which normalized repair benefits to the entire inspected fleet, suggested that CO had been reduced by eight percent and HC by six percent, with NOx increasing by one percent (Office of the State Auditor, 1999). While the study results seem contradictory, they cover different timeframes, make divergent assumptions (about deterioration rates, the fate of vehicles with final failures) and employ different measures in estimating I/M effectiveness.

One weakness in attributing before-after emission differences to I/M is the potential for "regression to the mean" emissions behavior, in which a portion of I/M failures will

register lower emissions on the final inspection without repair.[100] This phenomenon is driven by tremendous emissions test-to-test variability, the presence of vehicles with marginally failing emissions, and variance in environmental conditions favorable to test performance. Without verifying repairs, the emissions differences between initial and final test scores may overestimate program effectiveness.

E.2.2 Roadside Emission Inspections

Used primarily in California, roadside emissions tests are administered with the aid of law enforcement officers who randomly pull vehicles over and ask motorists to voluntarily submit their vehicles to an emissions inspection. Volunteer license plate numbers are then used to query the I/M program database to determine those vehicles with and without an inspection in the past twelve months. Recently inspected and uninspected vehicle emissions are then compared to estimate the emission reductions due to enhanced I/M. Roadside emissions estimates of 1999 enhanced I/M program effectiveness indicate emission reductions of 13 percent for CO, 14 percent for HC, and 6 percent for NO (California Air Resources Board, 2000).

---

[100] Regression to the mean occurs when two imperfectly correlated measures are compared for a nonrandom sample. The nonrandom sample is typically drawn from high or low scorers on either measure. Regression to the mean occurs when the sample mean moves towards the population mean in the absence of intervention. In the context of I/M evaluation, this means that certain vehicles failing their initial I/M test will score more closely to the mean of the population on the retest, i.e., register passing emissions, without repair. Regression-to-the mean can also occur in vehicles that pass their initial inspection but would fail a subsequent retest.

As with I/M program data, roadside pullovers enable the collection of odometer data for mileage estimates. In contrast with I/M program data, the spontaneity of roadside inspections preclude fraudulent test results that overestimate effectiveness, as well as pre-inspection maintenance behavior that underestimates program effectiveness. However, because roadside emissions tests employ a portable version of official inspection procedures, they sacrifice real-world driving conditions. Furthermore, the approach is costly and generates limited data, requiring as many as four technicians and one law enforcement officer to measure approximately 25 vehicles per day (Wenzel, et al 2000, p. III-8). Self-selection bias is a risk because the test is voluntary and tends to yield a ten percent refusal rate (Wenzel, et al 2000, p. III-8).[101]

E.2.3 Remote Sensing Data from Onroad Vehicles

A second source of data for evaluating I/M program effectiveness, the one used in this study, is from remote sensing devices (RSD) that measure the emissions of vehicles while they are being driven. The advantage of in-transit measurement is the ability to observe a vehicle's emissions under typical driving conditions, which cannot be as easily captured by traditional controlled emissions testing procedures. Remote sensors can measure a large number of vehicles, an important attribute given the need to control for

[101] The evidence of such bias is mixed. One recent study that used remote sensing to measure the vehicle emissions of refusals and participants alike found no significant difference between the two groups (Wenzel et al, 2000, pg. III-8), while an earlier similar study found that refusal vehicles had 2.5 times the emissions of volunteer vehicles (Stedman, 1994).

tremendous emissions variability due to vehicle type, age, make and model, and emission control technology. A final advantage stems from the unscheduled nature of the measurement, which precludes pre-inspection and fraudulent maintenance behavior that can occur when motorists (as with I/M tests) know when a measurement will occur.

In contrast with the highly controlled parameters of the emissions inspection, the physical circumstances of remote sensing data collection are only approximated through sampling site characteristics (e.g., moderate grades to ensure vehicles operate under only a slight engine load and sampling sites that avoid residential areas to minimize inflated emissions from cold engines). Another drawback is that remote sensors capture a split-second emissions reading that may not reflect a vehicle's typical emissions, making larger samples sizes a requirement to average out random emission fluctuations and to profile emissions aggregated within vehicle type (cars vs. trucks) and model year.

Remote sensing data has been used in three ways to evaluate I/M programs. The first method averages the emissions of vehicles measured before initial and after final I/M testing, with the difference attributed to I/M program effectiveness. Dubbed the "comprehensive method" in recent EPA evaluation guidance, emissions differences can also be generated for various subfleets, such as vehicles initially failing and ultimately passing I/M testing versus failing vehicles that never receiving a final pass. This approach enables a variety of I/M-related analyses, such as deterioration rates of I/M repairs, the influence of pre-I/M repairs on emissions baselines, and a comparison with estimates based on I/M records alone. The major disadvantage to this approach is the enormous volume of onroad data required to measure a representative sample of vehicles before and after I/M

testing. Sample size requirements hinge on the probability of measuring vehicles onroad within a specific time period of I/M testing, a probability that fluctuates with testing frequency and the distribution of sampling throughout the year.

The comprehensive method was used to estimate the effectiveness of the California South Coast Air Basin's enhanced I/M program in 1999 (Wenzel et al., 2000). "Smog Check" I/M records were used to delineate tested from untested vehicles by the existence of an enhanced inspection within the past twelve months.[102] A comparison of these vehicle groups indicates a ten percent reduction in CO, a four percent reduction in HC, and a five percent increase in NOx. An earlier remote sensing study in California in 1996 compared the onroad emissions of 3.5 million vehicles 30 to 90 days before with up to 90 days after their basic I/M test (Klausmeier and Weyn, 1997). For those vehicles that failed their initial smog check and then passed, both CO and HC emission differences registered at 20 percent. Normalizing this result to the entire fleet yielded an estimated nine percent emissions reduction in HC and CO. A third evaluation, of the Arizona enhanced I/M program in 1997, analyzed four million remote sensing measurements on 1.2 million vehicles in the Phoenix I/M area (Wenzel, 1999). The results indicated a seven percent reduction in CO and an 11 percent reduction in HC.

---

[102] Untested vehicles may have been inspected under the previous basic I/M program more than twelve months ago or they may have had an enhanced inspection after the remote sensing reading.

One weakness in the comprehensive method is the potential seasonal effects that results from the year-round testing required to obtain adequately sized samples. Users of this method have also tended to rely on a few high-volume sites, yielding a large number of repeat vehicles that lower the fraction of unique vehicles that could be reached at a greater number of sites.

A second I/M evaluation approach using remote sensing, known as the Step Method, compares inspected with uninspected vehicles during the first year of a new or upgraded program. The uninspected vehicles comprise an internal control group against which to compare the emission reductions of the inspected vehicles. Because the method applies to the early phases of a new or improved program, it can be used only once to assess program effectiveness.

A remote sensing study of the Colorado Enhanced I/M program compared odd (inspected) and even (uninspected) model year vehicles during the end of the first year of a new biennial enhanced I/M program (Stedman, et al, 1997). At that point, in program history, all odd model year vehicles should have been inspected, whereas all even model year vehicles had no reason to be inspected. This timing rendered even model year vehicles the untested control group against which to compare the odd model year vehicle emissions. The comparison of odd and even model year emissions suggested that Colorado's enhanced I/M program had reduced CO between five and nine percent, while HC and NO showed no improvement.

Three factors limit the generalizability of the Colorado study results to enhanced I/M program effectiveness. Remote sensing took place in a single location, which avoids any confounding socioeconomic or physical influences at different sites but limits generalizability to the overall fleet. Furthermore, vehicles traveling past the remote sensing site were decelerating, which does not represent typical driving conditions and is not the optimal condition for measuring carbon monoxide (Environ, 1998, p. 2-19). A third limitation was that the study measured vehicles transitioning from an annual basic I/M program to an enhanced I/M program, rendering it an evaluation of incremental program effectiveness and not a complete estimate of enhanced I/M program performance.

A third approach using remote sensing data (the one used in this study) compares the onroad emissions of vehicles registered in an I/M area to that of vehicles registered in non-I/M areas. The non-I/M area serves as a surrogate untested fleet. The validity of this approach relies on the selection of a non-I/M area comparable in fleet age, a well-documented contributor to vehicle emissions; climate, which can accelerate emission control equipment deterioration; and demographics, which influences the age, quality, and maintenance of the vehicle fleet. This approach was originally applied to the basic I/M program operating in four counties of the thirteen-county Atlanta ozone nonattainment area, with the nine nonattainment counties without I/M comprising the untested fleet. The analysis indicates that car and truck emissions for CO were 15 and ten percent higher, respectively, in the uninspected nine-county fleet than in the inspected four-county basic I/M fleet. The study is limited by its inability to control for differences in mileage and socioeconomic conditions between the two vehicle fleets.

E.3 I/M Program Evaluation components

This study employs an I/M program evaluation method that compares the onroad emissions differences observed in inspected and uninspected vehicles with the same emissions differences predicted by a U.S. Environmental Protection Agency mobile emissions model. The model-predicted emissions difference represents the goal of the I/M program, a reasonable assumption given that states use the model to generate the emission reduction credit received for automobile emissions testing programs. The emissions difference observed in onroad inspected and uninspected vehicles is assumed to reflect I/M program performance, an assumption rendered plausible only by the comparability of the inspected and uninspected fleets. We will devote attention in the next section to answering the comparability question.

This section describes the collection of data used in the evaluation. It details the Continuous Atlanta Fleet Evaluation (CAFE), the remote sensing study of onroad Georgia vehicles that provides onroad emissions data of inspected and uninspected vehicles. The MOBILE6, EPA's recommended emissions model, from which we extracted predicted emission factors, is also discussed. The last section outlines the algorithm that combines data from CAFE and MOBILE6 to generate effectiveness estimates for the Atlanta enhanced I/M program.

E.3.1 Onroad Emissions Data

The Continuous Atlanta Fleet Evaluation (CAFE) provides the onroad emissions data used to represent, *inter alia*, Atlanta enhanced I/M program performance. CAFE uses

remote sensing devices to measure annually the emissions of approximately 360,000 in-use vehicles in the 13-county I/M program area, as well as two cities located more than 75 miles from Atlanta that do not require vehicle emissions testing.[103] The study is an ongoing effort started in 1993 to collect vehicle emissions data for assessing a variety of trends, including fleet turnover, emission control deterioration, and socioeconomic impacts of mobile source control strategies.

RSD measures the emissions of passing vehicles remotely and unobtrusively so motorists are minimally aware of the equipment and do not alter their natural driving behavior. To that end, the remote sensing instrumentation is housed in a van parked on the roadside along with a videocamera. An infrared light source and its generator are placed on the opposite side of the road or on the median to create a beam of light that traverses the road. When a passing vehicle breaks the beam, it triggers a measurement of hydrocarbons, carbon monoxide, and nitrogen oxides in the exhaust. Simultaneously, a videocamera records the vehicle's license plate, which is automatically scanned into the database of emissions measurements.

After data collection, remote sensing measurements are merged with vehicle registration records using the vehicle license plate. The resulting database allows various characteristics of measured vehicles to be identified, including vehicle identification

[103] Augusta is located 136 miles east of Atlanta, whereas Macon is 76 miles south of Atlanta.

number,[104] make, model year, and vehicle type. License plates are also linked with inspection/maintenance records to identify vehicles with prior emission inspections.

RSD sampling sites are selected to ensure physically consistent but demographically diverse characteristics. Single straight lines of traffic with an average 35 mile-per-hour velocity are sought to facilitate single vehicle measurements and speeds that maximize measurement opportunities. Driver behavior and driving maneuvers are also observed at each site to ensure that remote sensing measurements would not be biased high by acceleration or low by coasting. Finally, notations are made during the site visits regarding any obvious or suspected diurnal patterns that exist which affect the traffic volume. If distinct variations are found to exist in sites ultimately selected, sampling times are scheduled to account for those diurnal patterns. U.S. Census tract data and traffic count reports inform the selection of different income ranges and land uses.

The remote sensing sites relevant to this study reside within the 13-county Atlanta I/M program area, 8 Atlanta counties without I/M program as well as the Georgia cities of Augusta and Macon. The latter locales do not require emissions testing and thus provided an uninspected vehicle fleet to serve as a control group for our previous I/M evaluations. These cities were chosen after a review of census data and registration records revealed

[104] Vehicle identification numbers are 17-digit alphanumeric strings that uniquely identify every vehicle manufactured. When decoded, they provide additional characteristics on vehicles. The VIN-decoded data of particular relevance to this research are vehicle type (car, truck, multi-purpose vehicle, van) and model year.

them to have characteristics – median household income, population density, and fleet distribution -- most similar to Atlanta than three other Georgia cities considered. But for the reasons which will be explained latter for the present analysis we also used as the reference point the data collected from the vehicles registered in 12 counties that surround Atlanta I/M program area.

According to the state regulation, effective April 1, 1999 the sulfur content of all gasoline supplied in a 25 Atlanta region[105] shall not exceed a seasonal average of 150 ppm (by weight) and, effective April 1, 2001, a per-gallon cap of 500 ppm (by weight)[106]. This rule made vehicle operational conditions in Atlanta 13-county nonattainment area and Augusta-Macon significantly unequal. Since there is no mechanism to separate benefits received from the usage of low sulfur gasoline and emission reductions due to I/M program, the usage of Augusta-Macon fleet as a control group for I/M program evaluation became questionable. In the effort to eliminate the fuel effect, the data collected from the vehicles registered in twelve counties that are not subject to the I/M program but receive the same fuel as Atlanta 13-counties I/M program area have been used in the present analysis as a reference point. The data collected on Augusta-Macon sites represents combined benefits form both I/M and GA fuel programs.

---

[105] 25-county Atlanta region include 13-county I/M program area and 12 additional counties without I/M program: Barrow, Bartow, Butts, Carroll, Dawson, Hall, Haralson, Jackson, Newton, Pickens, Spalding, Walton.

[106] Rules for Air Quality Control Chapter 391-3-1, July 20, 2005. http://www.gaepd.org/Files_PDF/rules/rules_exist/391-3-1.pdf

E.3.2 Predicted Emission Factors

We used MOBILE6, an EPA's recommended computer model for estimation of mobile emission factors, to predict emissions differences in inspected and uninspected vehicles.

E.3.3 Evaluation Algorithm

We estimated Atlanta enhanced I/M program effectiveness by comparing EPA model-predicted emission differences with observed emission differences in inspected and uninspected vehicles. The comparison yields a percentage that represents the proportion of expected emission reductions actually achieved by the program. The formula for estimating I/M effectiveness is as follows:

$$\text{Effectiveness}=\frac{\sum_{ij}[(O_{n_{ij}}-O_{m_{ij}})/O_{n_{ij}}](P_{n_{ij}})(C_{ij})(VMT_{ij})}{\sum_{ij}(P_{n_{ij}}-P_{m_{ij}})(C_{ij})(VMT_{ij})}$$

where: O_m and O_n are the average onroad emissions observed for a particular model year and vehicle type for I/M program and non-program vehicles, respectively; P_m and P_n are the model-estimated emission factors for I/M program and non-program vehicles for a given model year a vehicle type; C_{ij} is the fraction of the Atlanta fleet of that model year and vehicle type observed by CAFE; and VMT_{ij} is the average annual vehicle-miles-traveled by model year and vehicle type in the I/M program area.

The formula normalizes predicted and observed emissions differences in I/M program and non-I/M program vehicles by model year to the onroad fleet fraction and average annual mileage of that model year. This exercise enables the different units of measurement between onroad and predicted emissions – exhaust CO percentage/NOx ppm versus grams per mile of CO/NOx - to be put in ratio form.

E.4 Analysis

This section reports the results of the reference method for evaluating the Atlanta enhanced I/M program during its third two years of operation. The evaluation uses remote sensing emissions data collected in 2002 and emission factors predicted for the 2002 fleet by an EPA computer model. The 2002 calendar year represents the end of the third full cycle of enhanced IM testing and by this time all vehicles should have been inspected under new annual program.

Because the reference method involves direct comparisons between onroad data and EPA's MOBILE model, we restrict the data in several ways to obtain an "apples-to-apples" comparison. First, only 1978 to 2000 model year cars and 1980 to 2000 trucks are included in the analysis. While the Atlanta enhanced I/M program inspected back to the 1975 model year during 2002 low sample sizes among onroad vehicles precluded us from valid statistical analysis on cars manufactured earlier than in 1978. The 2001 and 2002 model years are not included since these vehicles were exempt from testing in 2002.

The second data restriction is the use of only vehicles registered in the thirteen Atlanta counties of the I/M program area as an *Inspected* fleet while some of the inspected vehicles could move to non I/M areas due to natural migration (such as change of ownership, etc.).

E.4.1 Data Overview

The remote sensing data used by this evaluation were collected at twenty-nine (29) Atlanta I/M program area sites and three non-program area sites in Augusta and Macon.[107] Measurements in the I/M program area were conducted from January to December 2002, while non-program area readings were collected over four days in March, September and December.

CAFE collected 209,436 measurements from vehicles registered in thirteen-county area with a 2002 inspection. In the non-program areas, 16,797 measurements were collected from vehicles registered in the counties comprising Augusta and Macon and 8,398 measurements from 12 Atlanta counties that are not subject to the I/M program. In both program and non-program areas, we randomly selected one measurement from unique vehicles with multiple readings.

This evaluation of the Atlanta enhanced I/M program relies on measurements of CO and NOx data. The primary reason for focusing on CO over HC is that the former pollutants

[107] I/M program area measurements are made within thirteen counties that comprise the metropolitan Atlanta ozone nonattainment area: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Fulton, Gwinnett, Henry, Forsyth, Paulding, and Rockdale. Non-I/M program measurements include Bibb, Richmond and Columbia counties.

have a smaller signal-to-noise ratio. CO's lower variance is due to its presence in higher concentrations in the atmosphere than HC, making it easier to measure by remote sensing devices and less susceptible to weather and driving conditions. In other words, although HC data was collected and analyzed during the study, we did not concentrate on it due to large reading errors.

E.4.2 Validity of Fleet Comparisons

Our ability to infer I/M effectiveness from the emission differences in Atlanta and Augusta/Macon vehicles hinges upon the comparability of the three fleets. The inspected Atlanta and uninspected Augusta-Macon vehicle fleets have similar model year distributions, although the inspected Atlanta fleet is slightly newer than the uninspected Augusta-Macon fleet (Figure E.1).

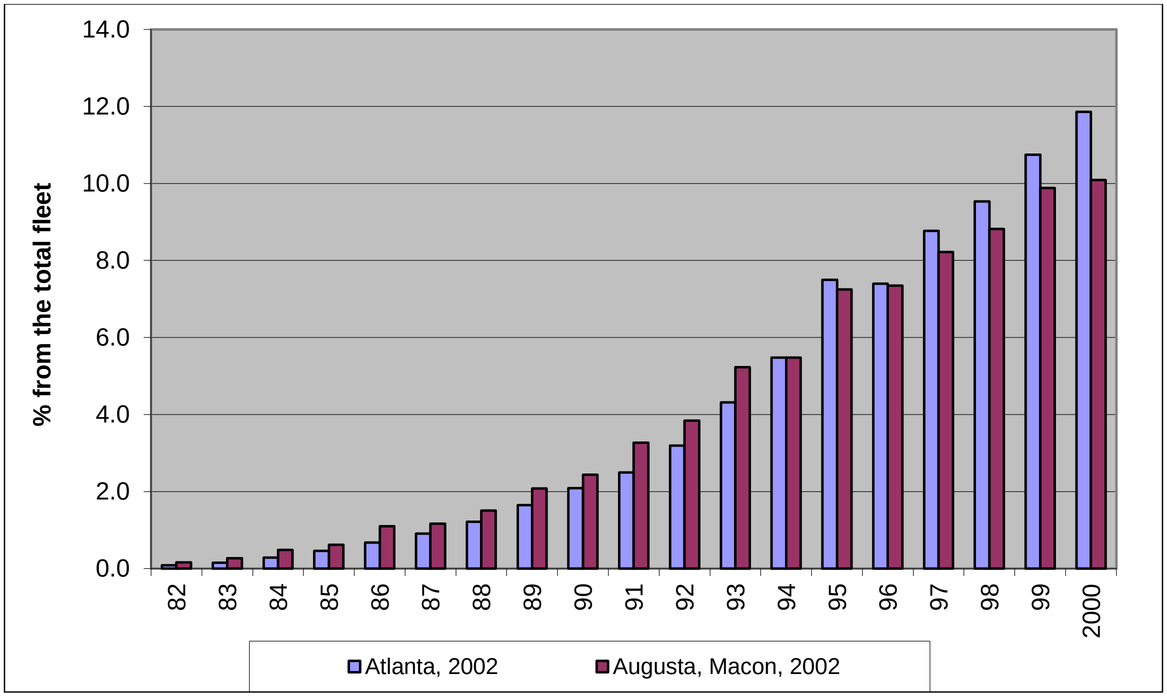

Figure E.1 Model Year Distributions of Inspected Atlanta Fleet and Uninspected Augusta-Macon Fleet.

A second issue for the validity of our analysis is whether the Augusta and Macon fleets are similar enough to be combined into one uninspected fleet. The previous studies had shown that the average CO emissions by model year and vehicle type do not differ significantly between the two fleets however, as was stated above, introduction of low sulfur gasoline initiative in 25 Atlanta counties changed this. Vehicles registered in Atlanta 13-county area seize benefits from two different emission control programs while vehicles from Augusta and Macon have neither. Taking into consideration former facts Augusta-Macon fleet was used to analyze a combined effect of I/M and fuel programs and data collected from 12 Atlanta counties that are not subject to I/M testing was used for estimation of I/M contribution.

E.4.3 Reference Method Results

The results of the reference method for evaluating the effectiveness of the Atlanta enhanced I/M program are laid out in Table E.1. But first, let us review the methodology for generating the estimates. We calculate the emissions difference in inspected and uninspected cars and trucks by model year and then weight those differences to that model year's annual average mileage and fleet fraction. The exercise is undertaken separately for predicted emissions factors and onroad emissions data. The weighted emissions differences in each category are then summed over all model years. The weighted value based on onroad emissions data becomes the numerator, whereas the weight value based on predicted emission factors becomes the denominator. Dividing the numerator by the denominator yields the percentage of expected emissions differences actually achieved in inspected and uninspected vehicles. The results of this exercise indicate that the effect of Atlanta enhanced I/M program captures 166 percent of CO reduction for cars and 229 percent for trucks as from predicted by EPA.

Table E.1 Effectiveness of Atlanta I/M Program and Fuel Program.

	Atlanta 13-counties inspected fleet vs. Augusta-Macon uninspected fleet	
	Cars	**Trucks**
CO	166%	229%
NOx, lower estimate	78%	68%
NOx, higher estimate	170%	150%

Delving into the data comprising these results, Figure E.2 and Figure E.3 compare the CO emissions differences in inspected thirteen-county Atlanta and uninspected Augusta-Macon vehicles measured onroad by RSD.

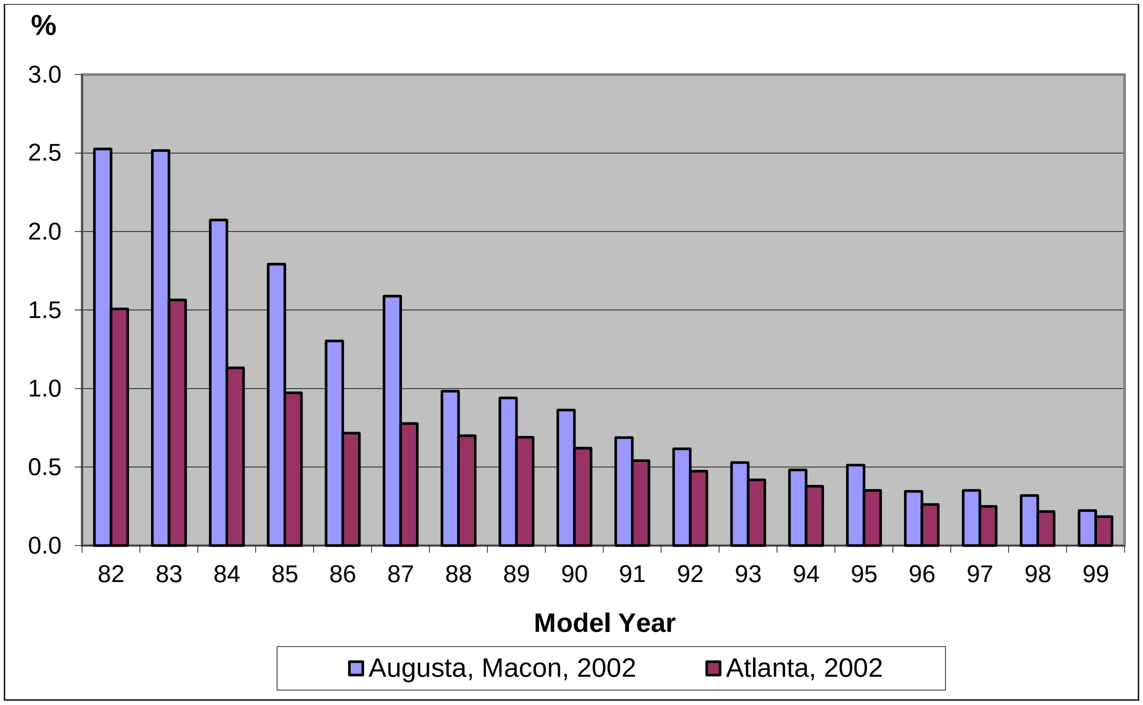

Figure E.2 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Cars Only.

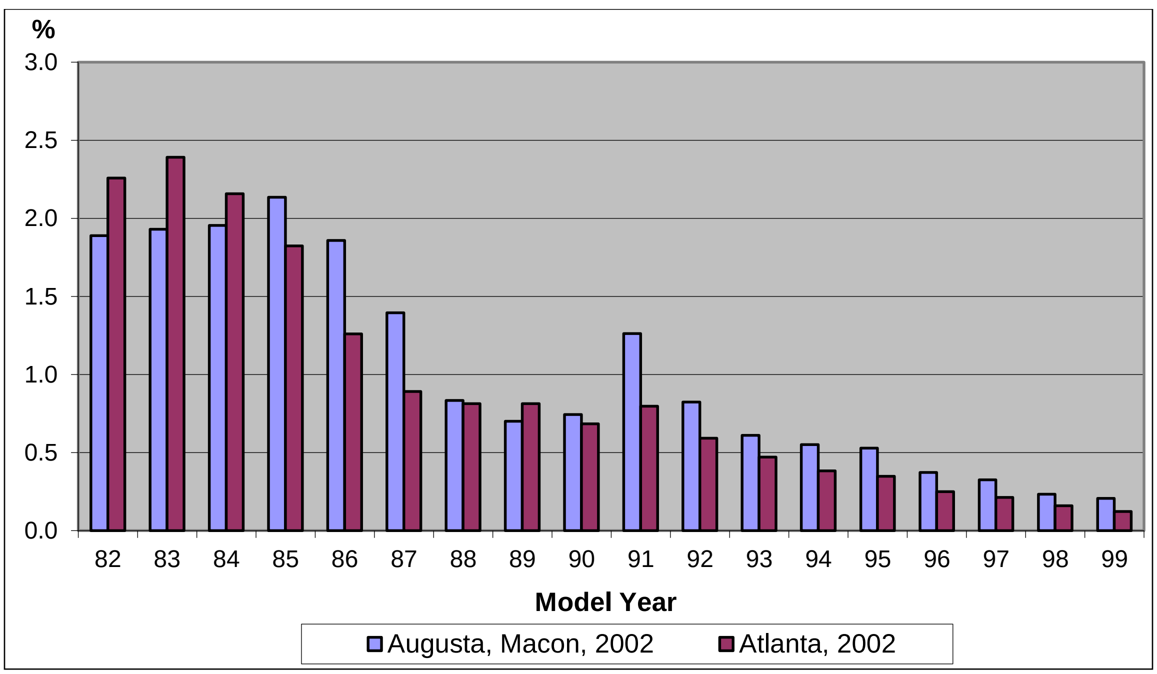

Figure E.3 Mean CO Comparison by Model Year for Atlanta Inspected Fleet and Augusta-Macon Noninspected Fleet. Trucks Only.

The onroad emission differences for NOx mimic this pattern, although with much larger fluctuations due to additional benefits Atlanta gets from the usage of low sulfur fuel. It is known that the amount of sulfur in the gasoline affects level of NOx exhausted. Figure E.4 illustrates the changes in the average NOx values due to seasonal variations of sulfur level in the gasoline supplied. Therefore additional references are needed to separate I/M air quality benefits and those from low sulfur fuel.

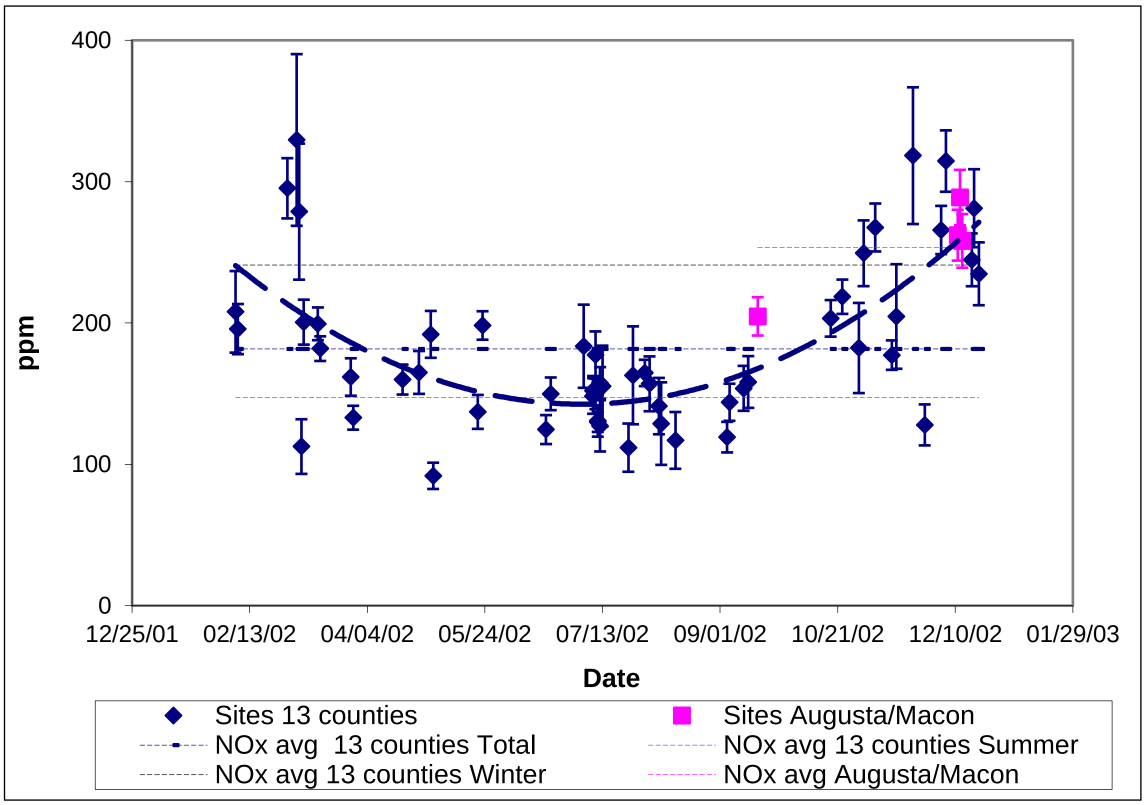

Figure E.4 Seasonal Average Values of NOx. Passenger Cars Only.

The data collected from vehicles registered in twelve Atlanta MSA counties that are not subject to I/M but operate on the same fuel was used as one of such references and represents the lower point of our estimated I/M benefits for NOx. The usage of Atlanta uninspected fleet as the only mark for comparison does not seems justified. First, all vehicles registered in Atlanta non I/M counties were measured on the significant distance from home location and may not represent the typical fleet as previous studies reveal that visiting vehicles are usually cleaner than domestic ones. Second, our estimation show that due to local migration between counties about 28% of vehicles that are not subject for I/M actually undergone the test during past two years. Therefore by comparing Atlanta inspected and uninspeced fleets we will underestimate benefits from I/M program.

For the reasons discussed earlier in this document the NOx data collected from Augusta and Macon can not be used for I/M evaluation directly. However, after several additional assumptions and modifications it can be utilized as the higher point of estimated I/M benefits. Our analysis has shown that measured CO levels are very little or not at all affected by the usage of low sulfur fuel or seasonality. Figure E.5 represents the CO curves for Atlanta 13 counties nonattainment area and Augusta-Macon

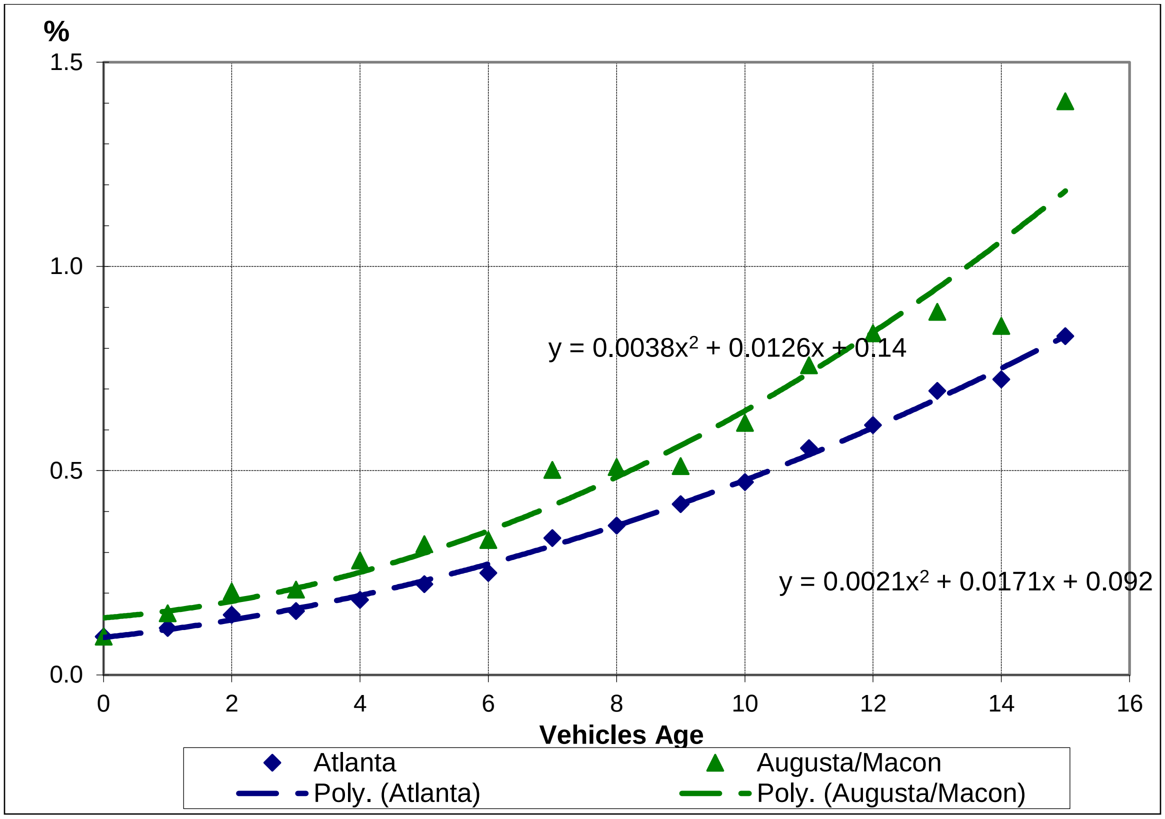

Figure E.5 Average CO Levels. Atlanta 13 Counties Area and Augusta-Macon. Passenger Vehicles Only.

I/M benefits can be represented as the difference of two integers from the functions that describe those curves: $\int_{2}^{15}(a_2x^2 + b_2x + c_2)dx - \int_{2}^{15}(a_1x^2 + b_1x + c_1)dx$.

On the other hand the ratio: $\dfrac{\int_{2}^{15}(a_2x^2 + b_2x + c_2)dx}{\int_{2}^{15}(a_1x^2 + b_1x + c_1)dx}$ gives us the universal coefficient by which emission levels measured in Atlanta I/M area differ from emission levels measured in Augusta and Macon. This exercise produced the coefficient that equals 2.2

and the higher estimation of I/M benefits for NOx can be found by multiplying the lower point of I/M benefits values on this coefficient.

E.5 Discussion

Interpreting emissions differences in the Atlanta inspected fleet and Augusta/Macon fleets as combined effect of the enhanced I/M program and fuel programs assumes that we have controlled for all differences in these fleets. This assumption is challenged by the possibility that the Augusta/Macon fleet is composed of higher mileage or poorer quality vehicles than the Atlanta thirteen-county fleet. One source of evidence for mileage differences, the U.S. Department of Transportation data on daily vehicle miles traveled (VMT), suggests that vehicles in Atlanta travel 34 miles per day per capita versus 22 miles for vehicles in Augusta. This information would seem to weaken any hypothesized mileage difference, at least between Augusta and Atlanta. However, because GDOT estimates are based on observed freeway traffic flows that capture out-of-state as well as local vehicles, it is difficult to extrapolate these VMT estimates to the local vehicle fleet. Exclusion of luxury cars from analyzed data sets did not make significant changes in emission patterns therefore the fleet composition differences between Atlanta and Augusta/Macon are negligible.

The comparison of Atlanta thirteen counties inspected fleet with Atlanta uninspected fleet has the same validity issues. Since vehicles that likely to fail testing have the tendency for migration in neighboring counties that are not subject for I/M program we may overestimate its effectiveness. But due to close proximity inspected vehicles also penetrate the noninspected area after change of ownership or under other circumstances which leads to underestimation of I/M benefits.

E.6 Comparing results with previous review.

The reference method for evaluating vehicle inspection/maintenance programs yields several advantages over other methods using onroad remote sensing data. In fact, the reference method could be repeated over time to measure incremental effectiveness as more of the fleet is tested, inspectors become adept at identifying noncompliant vehicles, repair technicians gain experience at repairing emission control failures, and (more pessimistically) motorists learn better how to co-opt the test.

The study presented evaluates the third two-year period of the established in Atlanta thirteen counties IM program. The first evaluation review covered the 1997-1998 years and the second evaluation covered 1999_2000 years. Table E.2 summarizes results yielded by all three reviews.

Table E.2 IM Effectiveness Estimated for 1998, 2000 and 2002 Measurement Year.

Estimated IM Effectiveness		Cars	Light Trucks
1998 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	87%	75%
	NOx	NA	NA
2000 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	84%	84%
	NOx	NA	NA
2002 Measurement year Atlanta inspected fleet vs. Augusta/Macon fleet	**CO**	166%	229%
	NOx Low estimation	78%	68%
	NOx High estimation	170%	150%

Unfortunately for the reasons discussed in previous sections of this report results seized by previous and current analyses can not be directly compared. In current study I/M effectiveness represented by figures derived not only from comparison of inspected Atlanta and uninspected Augusta/Macon fleets but also by using uninspected Atlanta 12 counties fleet as an additional reference point. The comparison of vehicles registered in Atlanta I/M program area and those from Augusta/Macon represents the combined effect of I/M and state fuel programs wile evaluation of inspected and uninspected Atlanta fleets correspond only to benefits from I/M program.

As for the Atlanta enhanced I/M program overall, the reference method suggests that the program is reducing onroad emissions, but may not be meeting EPA model predictions for NOx emissions. Future research efforts will include replicating the reference method at different points in place and time to estimate the impact of emission reduction programs changes discussed in this study on onroad fleet emissions.

The reference method is not without its limitations, however. Selecting a comparable non-program fleet is a challenging task, to say the least. While differences in fleet age and car/truck composition are relatively easy to account for between I/M and non-I/M fleets, discrepancies in maintenance trends, socioeconomic conditions and vehicle quality are difficult to discern. Until additional studies of non-I/M fleets shed light on the role of these influences on fleet emissions, equating fleet differences with I/M effectiveness will be a tentative proposition.

BIBLIOGRAPHY

Ando, A., McConnell, V., & Harrington, W. (1999). *Costs, Emission Reductions, and Vehicle Repair: Evidence from Arizona.* Washington, D.C.: Resources for the Future.

Bryner, G. C. (1995). *Blue Skies, Green Politics.* Washington, D.C.:Congressional Quaterly, Inc.

Cadle, S. H., Gorse, R. A., Belian, T. C., & Lawson, D. R. (1999). Real-World Vehicle Emissions: A Summary of the Eight Coordinating Research Council On-Road Vehicle Emissions Workshop. *Journal of the Air and Waste Management Association, 49*, 242-255.

California Air Resource Board and Bureau of Automotive Repair. (2010). *Evaluation of the California Smog Check Program Using Random Roadside Data.*

California Air Resources Board. (2000). *Evaluation of California's Enhanced Vehicle Inspection and Maintenance Program (Smog Check II).*

Choi, D., Beardsley, M., Brzezinski, D., Koupal, J., & Warila, J. (2010). MOVES sensitivity analysis: the impacts of temperature and humidity on emissions. *US EPA–Proceedings from the 19th Annual International Emission Inventory Conference.* Ann Arbor, MI: US EPA.

Committee on Vehicle Inspection/Maintenance Programs, National Research Council. (2001). *Evaluating Vehicle Inspection/Maintenance Programs* . Washington, D.C.: National Academy Press.

Corley, E., DeHart-Davis, L., Lindner, J., & Rodgers, M. O. (2003). Inspection/Maintenance Program Evaluation: Replicating the Denver Step Method for an Atlanta Fleet. *Environmental Science and Technology, 37*(12), 2801-2806.

DeHart-Davis, L., Corley, E., & Rodgers, M. (2002). Evaluating Vehicle Inspection/Maintenance Programs Using Onroad Emissions Data: The Atlanta Reference Method. *Evaluation Review, 26*(2), 111-146.

Environ, Inc. (2003). *Performance Audit of the Colorado Automobile Inspection and Readjustment (Air) Program.* Air Sciences, Inc.

Glover, E., & Brzezinski, D. (1997). *Analysis of the Arizona IM240 Test Program and Comparison with the TECH5 Model.* U.S. Environmental Protection Agency, Air and Radiation, Office of Mobile Sources.

Harrington, W., McConnell, V., & Ando, A. (2000). Are Vehicle Emission Inspection Programs Living Up to Expectations? *5*, pp. 153-172. Transportation Research Part D-Transport and Environment.

Harrington, W., McConnell, W., & Cannon, M. (1998). *A Behavioral Analysis of EPA's MOBILE Emission Factor Model.* Washington, D.C.: Resources for the Future.

Klausmeier, R., & Weyn. (1997). *Using Remote Sensing Devices to Evaluate the California Smog Check Program: Report to the California Bureau of Automotive Repair*. California Bureu of Automotive Repair.

Lawson, D. R. (1993). "Passing the Test"-- Human Behavior and California's Smog Check Program. *Journal of the Air and Waste Management Association, 43*, 1567-1575.

Lawson, D. R. (1995). The Costs of "M" in I/M-- Reflections on Inspection/Maintenance Programs. *Journal of the Air and Waste Management Association, 45*, 465-476.

National Research Council. (2001). *Transforming Remote Sensing Data into Information and Applications.* Washington, DC: The National Academies Press. Retrieved from https://doi.org/10.17226/10257

Neter, J., Kutner, M., Nachtsheim, C., & Wasserman, W. (1996). *Applied Linear Statistical Models.* Chicago, IL: Times Mirror Higher Education Group.

Pierson, W. R. (1996). New Directions: Motor Vehicle Inspection and Maintenance Programs-- How Effective Are They? *Atmospheric Environment, 30*(21), i-iii.

Rodgers, M., DeHart-Davis, L., & Lorang, P. (1999). Measuring Inspection/Maintenance Program Effectiveness Using Optical Remote Sensing: Results of the Continuous Atlanta Fleet Evaluation. *Atmospheric Environment*.

Samoylov, A. (2013). *Improvement of the Efficiency of Vehicle Inspection and Maintenance Programs Through Incorporation of Vehicle Remote Sensing Data and Vehicle Characteristics.* Atlanta: GaTech.

Sierra Research. (1998). *Review of Air Quality and Motor Vehicle Technology Issues Pertaining to the Design of AirCare II.* Technical Report, Sacramento, CA.

Spindt, R., Dizak, R., Stewart, R., & Meyer, W. (1979). *Effect of Ambient Temperature on Vehicle Emissions and Prformance Factors.* Ann Arbor: US EPA.

State of Colorado Office of the State Auditor. (1999). *1999 Audit of the Colorado Air Improvement Resource Program.* Novi, Michigan: Air Improvement Resource, Inc.

Stedman, D. H. (1996, December). EPA Overestimates the Benefit of Biennial Inspection and Maintenance. *EM Feature*, 27-29.

Stedman, D. H., Bishop, G., & Slott, R. (1998). Repair Avoidance and Evaluating Inspection and Maintenance Programs. *Environmental Science & Technology, 32*(10), 1544-1545.

Stedman, D., Bishop, G., Aldrete, P., & Slott, R. (1997). On-Road Evaluation of an Automobile Emission Test Program. *Environmental Science & Technology, 31*(3), 927-931.

U.S. Environmental Protection Agency. (1978). *Guidance for Vehicle Inspection/Maintenance Programs.* Technical Report, U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. (2001). *Guidance on Use of Remote Sensing for Evaluation of I/M Program Performance.* U.S. Environmental Protection Agency.

Wenzel, T. (1998). Evaluation of Arizona's Enhanced IM Program. *Clean Air Conference.* Brechenridge, CO.

Wenzel, T. (1999). *Using Program Test Data to Evaluate the Phoenix I/M Program: Report to the Arizona Department of Environmental Quality, Lawrence Berkeley*

National Laboratory. Retrieved 2018, from Arizona Department of Environmental Quality: https://legacy.azdeq.gov/environ/air/vei/download/lawreport.pdf

Wenzel, T., Gumerman, E., Singer, B., & Sawyer, R. (2000). *Evaluation of the Enhanced Smog Check Program: A Report to the California Inspection and Maintenance Review Committee.* Retrieved from Fuel Efficiency Automobile Test Data Center: http://www.feat.biochem.du.edu/assets/reports/CA%20IMRC%20Report/IMRC20000619Part3Eval.PDF

Wenzel, T., Singer, B., & Slott, R. (2000). Some Issues in the Statistical Analysis of Vehicle Emissions. *Journal of Transportation and Statistics*, 1-14.

Wooldridge, J. M. (2000). *Introductory Econometrics.* South-Western College Publishing.

Younglove, T. S. (2005). Designing on-road vehicle test programs for the development of effective vehiclt emission models. *Transporation Research Board* (pp. 46-56). Washington: Transportation Research Record 1941.